

**LISSY VIANEY LIZARAZO ROJAS**

**MODELACIÓN ECONOMETRICA PARA EVALUAR LA SEVERIDAD DE  
ACCIDENTES EN CARRETERAS. APLICACIÓN A VÍAS INTERURBANAS  
COLOMBIANAS.**

**Proyecto de grado para optar por el título de.**

Magister en Ingeniería civil.

Director:

Ing. Víctor Cantillo Maza, PhD

Ing. Julián Arrellana, PhD

UNIVERSIDAD DEL NORTE

BARRANQUILLA, MAYO 2015

**LISSY VIANEY LIZARAZO ROJAS.**

Tesis presentada al comité conformado por:

---

---

---

**UNIVERSIDAD DEL NORTE**  
**BARRANQUILLA, MAYO 2015**

## **AGREDECIMIENTOS.**

Inicialmente, quiero agradecerle a Dios por que solo él sabe todo lo que he pasado para llegar aquí gracias papito dios por darme la fuerza y la sabiduría suficiente para cumplir unas más de mis metas.

A una persona muy importante para mí, CARLOS. Sin el nada de esto fuera posible. Gracias por soñar por mí, gracias por desear tantas cosas buenas para mí. Sin tu fuerza y tu apoyo nada de esto fuera realidad, gracias y mil gracias por esas palabras “puedes contar conmigo no hasta 1 ni hasta 10 si no contar conmigo”

A mi tutor de tesis VÍCTOR CANTILLO, porque fue como un padre que cuida a sus hijos: apoyándolos y creyendo en ellos a pesar de sus caídas. Gracias por estar ahí, por ayudarme en este proceso de formación, y por hacerme desear parecerme a usted cada día más.

Al ingeniero JULIÁN ARELLANA, quien con su ayuda incondicional me hizo sentir que mi tesis tomaba forma. Fue como la viga de cimentación, que sin ella no se puede construir una buena estructura.

A la persona que me ayudó a realizar mis primeros modelos, JANER GALVAN. Gracias por sus explicaciones y tiempo dedicado.

A mi querida amiga, ANDREA AREVALO. Desde el inicio que empezó esta travesía, desee compartirla contigo. Gracias por estar dispuesta a atravesar este camino y llegar a la meta juntas.

A mis compañeros de maestría: YOLI, LAURA, MAURICIO, DELIA, CLAUDIA, MAJO, LUIS, JULIO, JORGE, porque de cada uno aprendí cosas nuevas. Doy gracias a la vida que nos permitió conocernos en esta etapa de nuestras vidas.

Por último, y con igual importancia que Dios está mi familia. A mi PADRE y mi MADRE, gracias por estar siempre ahí y por darme todo su amor incondicional. Fue lo que me mantuvo llena de fuerza. A mis HERMANOS, cada uno aportó un granito para que mis sueños, que también son sus sueños, se hicieran realidad.

# TABLA DE CONTENIDO

1	resumen.....	7
2	INTRODUCCIÓN.....	8
2.1	Descripción del problema .....	8
2.2	Situación accidentalidad vial en Colombia.....	10
2.3	Objetivo general .....	14
2.4	Hipótesis .....	14
2.5	Contribuciones .....	15
2.6	Estructura de la tesis .....	15
3	ANTECEDENTES .....	16
3.1	Modelo Logit Multinomial MNL .....	21
3.2	Modelos de respuesta ordinal.....	22
3.3	Modelo Logit Mixto Con Componentes De Error. ....	24
4	DESCRIPCIÓN DE LOS DATOS .....	26
5.	DESCRIPCIÓN DE LOS MODELOS .....	38
5.1.	Modelos Estimados De La Severidad De Los Eventos Con La Base De Datos (2007 - 2009). 38	
5.2.	Modelos Estimados De La Severidad De La Lesión De La Víctima Con La Base De Datos (2011 -2012) .....	46
5	CONCLUSIONES.....	49

## TABLA DE FIGURA

<b>Figura 1 Tasa departamental por sexo por cada 100.000 habitantes, Muertes y lesiones en accidentes de tránsito .....</b>	<b>10</b>
<b>Figura 2. Muertes por accidentes de transporte. Colombia, 2004-2013 .....</b>	<b>11</b>
<b>Figura 3 Lesiones por accidentes de transporte. Colombia, 2004-2013.....</b>	<b>12</b>
<b>Figura 4 Gravedad de los accidentes – total nacional 2009 – 2012.....</b>	<b>13</b>
<b>Figura 5. Factores que influyen en la severidad de los accidentes de tránsito. ....</b>	<b>17</b>
<b>Figura 6 Modelos Utilizados en investigaciones anteriores para evaluar severidad de accidentes de tránsito. ....</b>	<b>19</b>
<b>Figura 7 Esquema del modelo ordinal. ....</b>	<b>24</b>
<b>Figura 8. Ilustración de los 11.500 km de vía colombiana a las cuales se les realizó el inventario vial .....</b>	<b>27</b>
<b>Figura 9 Distribución porcentual de los anchos de carril de los accidentes. ....</b>	<b>31</b>
<b>Figura 10 Distribución porcentual de los anchos de berma en los accidentes. ....</b>	<b>31</b>
<b>Figura 11. Distribución porcentual de la gravedad de los accidentes por año. ....</b>	<b>32</b>
<b>Figura 12 Distribución de las velocidades de operación de los accidentes. ....</b>	<b>32</b>
<b>Figura 13 Distribución de las velocidades de operación de los accidentes según la gravedad del accidente.....</b>	<b>33</b>
<b>Figura 14 Distribución de presencia de curvas en los accidentes. ....</b>	<b>33</b>
<b>Figura 15. Distribución de la severidad de accidente según el tipo de sexo.....</b>	<b>36</b>
<b>Figura 16 Distribución De La Severidad De Accidente Según El Estado Civil.....</b>	<b>36</b>
<b>Figura 17 Distribución De La Severidad De Accidente Según la característica de la víctima .....</b>	<b>37</b>

## LISTA DE TABLA

<b>Tabla 1 Accidentes por gravedad, según el área de ocurrencia – total nacional 2009- 2012.</b>	
.....	14
<b>Tabla 2 Características de los datos.....</b>	<b>29</b>
<b>Tabla 3. Estadísticas descriptivas de las variables base de datos 2007 - 2009.....</b>	<b>30</b>
<b>Tabla 4 Características y tipos de las variables a utilizar en los modelos para la base de datos 2011- 2012 .....</b>	<b>34</b>
<b>Tabla 5 Características de los datos.....</b>	<b>35</b>
<b>Tabla 6 Características y tipos de las variables a utilizar en los modelos para la base de datos 2011- 2012 .....</b>	<b>37</b>
<b>Tabla 7 Definición de las variables a utilizar en la modelación .....</b>	<b>40</b>
<b>Tabla 8 Modelo Multinomial Estimado .....</b>	<b>41</b>
<b>Tabla 9 Modelo Logit Ordinal .....</b>	<b>43</b>
<b>Tabla 10 Modelo Logit Mixto Con Componentes De Error .....</b>	<b>46</b>
<b>Tabla 11 Modelo Logit Multinomial.....</b>	<b>47</b>

# **1 RESUMEN.**

La presente tesis evalúa las variables que inciden en la severidad de los accidentes y en la severidad de las lesiones de las víctimas del accidente en vías interurbanas colombianas utilizando modelos de la familia logit, a través de sus estructuras multinomial, ordinal y mixto con componentes de error, en los cuales se han definido niveles de severidad asociados a la gravedad del accidente. Los datos de accidentes de las vías interurbanas utilizados en la investigación provienen de información geo-referenciada, que incluye características de la vía y el entorno (geometría, uso de suelo, presencia de obstáculos, visibilidad, estado del pavimento, entre otros), y el tránsito. Los resultados indican que los métodos econométricos propuestos son aproximaciones apropiadas para identificar los factores que determinan el grado de severidad en accidentes. Las variables más significativas son la sección transversal (ancho de carriles y bermas), la magnitud del tránsito. Los modelos estimados sugieren que la gravedad del accidente se incrementa cuando están involucrados peatones y motociclistas.

Los resultados de la estimación de los modelos para explicar la gravedad de accidentes en carreteras colombianas muestran mejores resultados cuando se especifican modelos que consideran el orden en el nivel de gravedad de los accidentes. Al realizar una comparación entre los modelos de componentes de error y el logit Multinomial se pudo concluir que la varianza del termino error, que captura la correlación entre alternativas, no resultó ser significativa por lo cual se prefiere la simplicidad del modelo logit multinomial reflejando un buen ajuste estadístico cuando se comparan sus resultados con los del modelo de componentes de error. Realizando un comparativo entre el modelo convencional logit multinomial y el modelo logit ordinal se concluyó que el modelo que mejor se ajusta a los datos es logit ordinal, porque el modelo logit ordinal tiene en cuenta la naturaleza de los datos, pues este último modelo reconoce que es más severo un accidente con muerto que un accidente con heridos y un accidente con heridos que un accidentes con solo daños materiales.

## **2 INTRODUCCIÓN**

La necesidad de minimizar las consecuencias de los accidentes de tránsito a través de la proyección de programas preventivos implica estudiar en detalle los factores que inciden en la gravedad de estos eventos. En Colombia, los accidentes de tránsito se han convertido en un problema de salud pública. Según estadísticas oficiales (Instituto Nacional de Medicina Legal y Ciencias Forenses, 2013) la accidentalidad vial en el 2013 dejó 6.219 personas muertas y 41.823 víctimas no fatales, dentro de las cuales el 81% de las muertes corresponden a personas de sexo masculino. Además, se ha identificado que los grupos más vulnerables resultaron ser los peatones y los usuarios de motocicleta.

La accidentalidad en Colombia ha sido un problema que ha venido creciendo en los últimos años a pesar de las campañas de prevención que se vienen practicando. El impacto que los accidentes de tránsito tienen en la sociedad es significativo. Por esta razón, es fundamental evaluar qué factores inciden en la severidad de los accidentes en las carreteras del país.

El Gobierno de Colombia, a través del Ministerio de Transporte, ha diseñado el Plan Nacional de Seguridad Vial (PNSV 2011 – 2016) el cual está estructurado fundamentalmente siguiendo las directrices de la Organización Mundial de la Salud (OMS), con el objetivo reducir en un 50% el número de muertos derivados de los accidentes de tránsito.

En esta investigación, se proponen modelos econométricos para evaluar los factores que inciden en la severidad de los accidentes en las carreteras colombianas, utilizando modelos de la familia logit (multinomial, mixto y ordinal). Los datos de accidentes de las vías interurbanas utilizados en la investigación provienen de información geo-referenciada, que incluye la severidad del accidente y las características de la vía y el entorno (geometría, flujo vehicular, el uso de suelo, presencia de obstáculos, visibilidad, estado del pavimento, entre otros).

### **2.1 Descripción del problema**

Cada año, en el mundo mueren cerca de 1,2 millones de personas por accidentes de tránsito y entre 20 y 50 millones de personas sufren traumatismos provocados por la misma causa según estimaciones realizadas por la OMS (2013). Los países con mayor tasa de muertes en accidentes



de tránsito son los de ingresos bajos y medios, en los cuales se presentan cerca del 90% de las muertes.

Según la OMS (2013) las lesiones causadas por accidentes de tránsito son la octava causa de muerte y primera causa de muerte entre edades de 15 a 29 años a nivel mundial. El reporte advierte que si no se adoptan medidas inmediatas y eficaces, en el año 2030 los traumatismos se convertirán en la quinta causa de muerte a nivel mundial. Ello se debe, en parte, al rápido aumento del mercado de vehículos de motor sin que haya mejoras suficientes en las estrategias sobre seguridad vial ni en la planificación del uso del territorio.

Las estadísticas disponibles en Colombia (Instituto Nacional de Medicina Legal y Ciencias Forenses, 2013) indican que los peatones y motociclistas constituyen el grupo más vulnerable dado que representan más del 50% de las muertes por accidentes de tránsito. Específicamente, los motociclistas ocupan el primer lugar con un 44,28% en muertes y 22,67% en heridos, seguido por el peatón el cual representa 29,27% en muertes y 22,67 % en heridos. En referencia a rangos etarios, es relevante señalar que los jóvenes entre 15 y 34 años constituyeron el 43,70% de las muertes registradas en el 2013.

El análisis geográfico de los índices de accidentalidad en el país muestra diferencias importantes entre regiones, tal como se aprecia en la Figura 1. Destaca el caso de 5 departamentos (Casanare, Arauca, Meta, Cesar y Huila) que concentran el 13,30% de las muertes siendo sólo el 7,80 % de la población nacional, teniendo los índices relativos a la población más desfavorables.

La finalidad del proyecto es identificar los factores de la infraestructura, del entorno, del tránsito que influyen en la severidad de los accidentes en las carreteras interurbanas de Colombia. Para ello se utilizarán herramientas econométricas que permitan explicar la severidad de los accidentes en función de los factores indicados.

La presente investigación busca contribuir a entender los factores que inciden en la severidad de accidentes en vías interurbanas. Esta información se considera indispensable para establecer políticas e intervenciones que contribuyan a disminuir las pérdidas sociales derivadas de la accidentalidad.

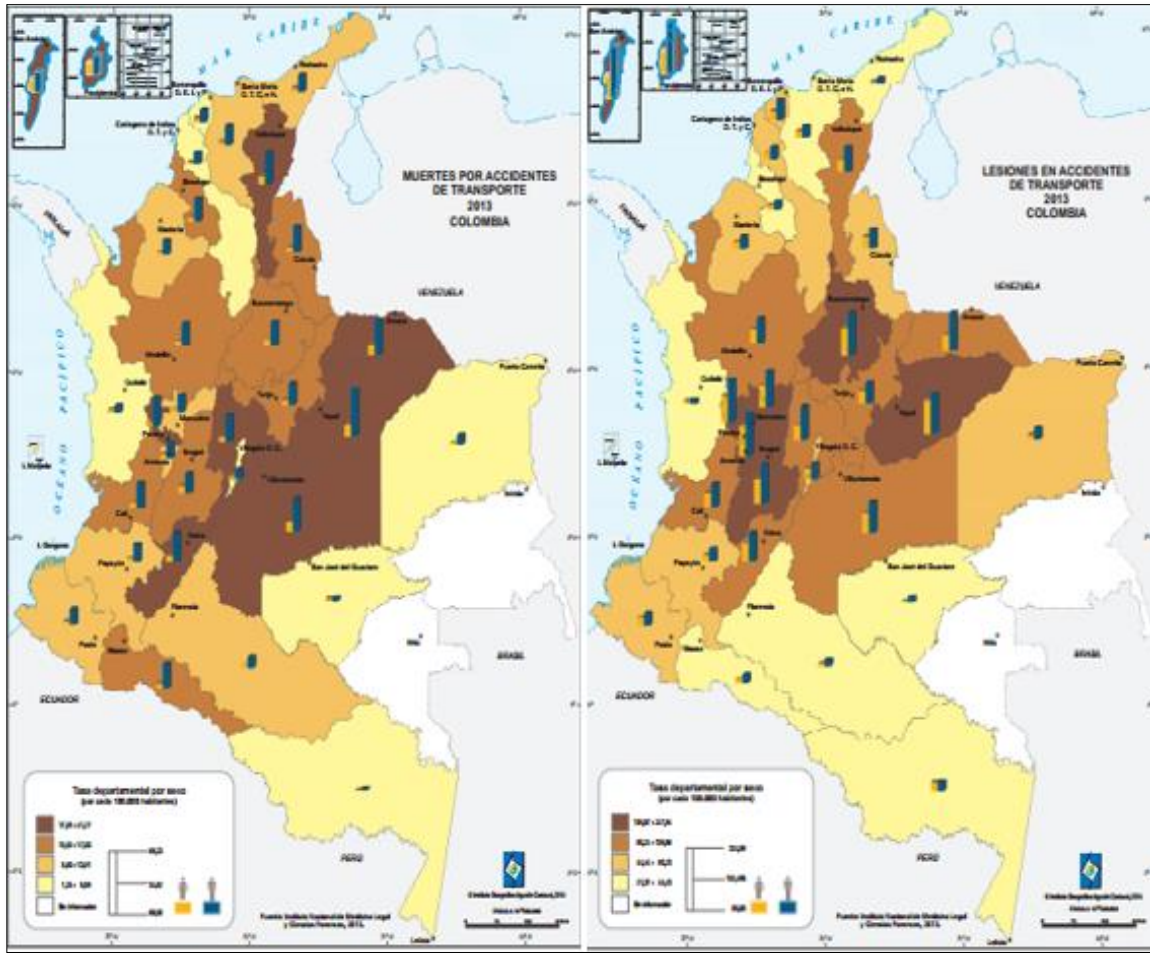


Figura 1 Tasa departamental por sexo por cada 100.000 habitantes, Muertes y lesiones en accidentes de tránsito  
Fuente: Instituto Nacional de Medicina Legal y Ciencias Forenses, 2013

La investigación busca responder las siguientes preguntas:

- ¿Cuáles son los factores de la infraestructura, del ambiente y el tránsito que determinan la severidad de los accidentes?
- ¿Cuál enfoque econométrico es el más apropiado para evaluar las causas que influyen en la severidad de los accidentes?

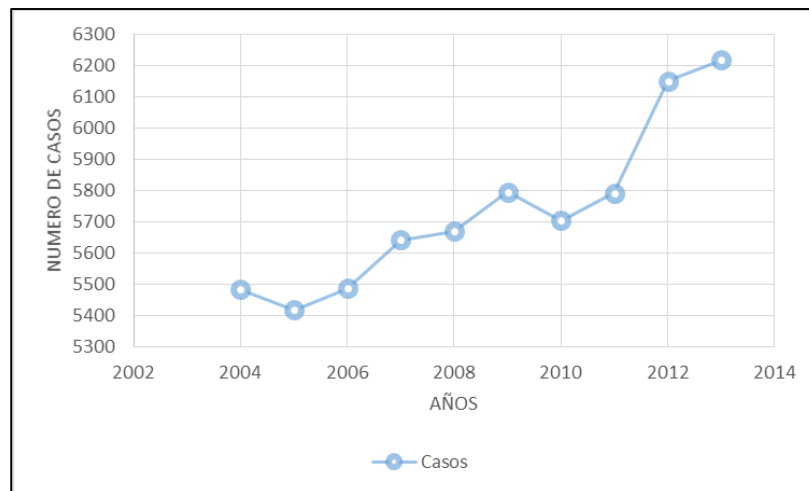
## 2.2 Situación accidentalidad vial en Colombia.

La recolección de antecedentes asociados a la siniestralidad en el tránsito es realizada por el Instituto Nacional de Medicina Legal y Ciencias Forenses, que considera los fallecidos hasta 30

días después de ocurrido el accidente, lo que es coincidente con el método internacional de los países que han desarrollado de manera sostenible políticas en seguridad vial.

Durante el 2013, el Instituto Nacional de Medicina Legal y Ciencias Forenses registró 48.042 casos atendidos por accidentes de transporte. Las lesiones fatales corresponden a un total de 6.219 personas fallecidas (12,94%) y las lesiones no fatales ascienden a un total de 41.823 personas lesionadas (87,06%). Se resalta que en 2013 ocurrió el más alto número de muertes de la última década<sup>1</sup>(Instituto Nacional de Medicina Legal y Ciencias Forenses, 2013).

En Colombia durante 2013 se tuvo una tasa de 13,2 muertes por cada cien personas. Aunque es inferior a la tasa registrada por la región de las Américas (16,1), preocupa que es la tasa más alta de la última década. La evolución de dicha tasa se muestra en la Figura 2.



**Figura 2. Muertes por accidentes de transporte. Colombia, 2004-2013**

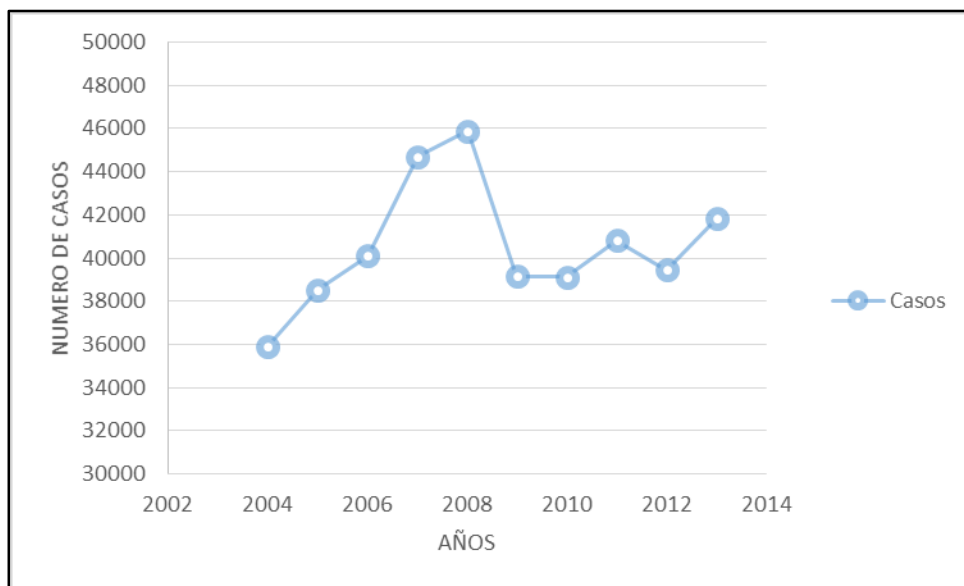
**Fuente:** Elaboración Propia a partir de datos de Instituto Nacional de Medicina Legal y Ciencias Forenses.

Por otra parte, en 2013 las víctimas no fatales (heridos) incrementaron 6,04% respecto al año anterior. En el 2013 el número de lesionados por accidentes de transporte informado al Instituto Nacional de Medicina Legal es el tercer registro más alto de la última década, por debajo tan solo de los años 2007 y 2008.

De acuerdo a lo registrado por los sistemas de información del Instituto Nacional de Medicina Legal, la tasa de lesionados por 100.000 habitantes no presenta variaciones significativas en la última década, exceptuando los valores atípicos registrados en el año 2007 y 2008 Figura 3.

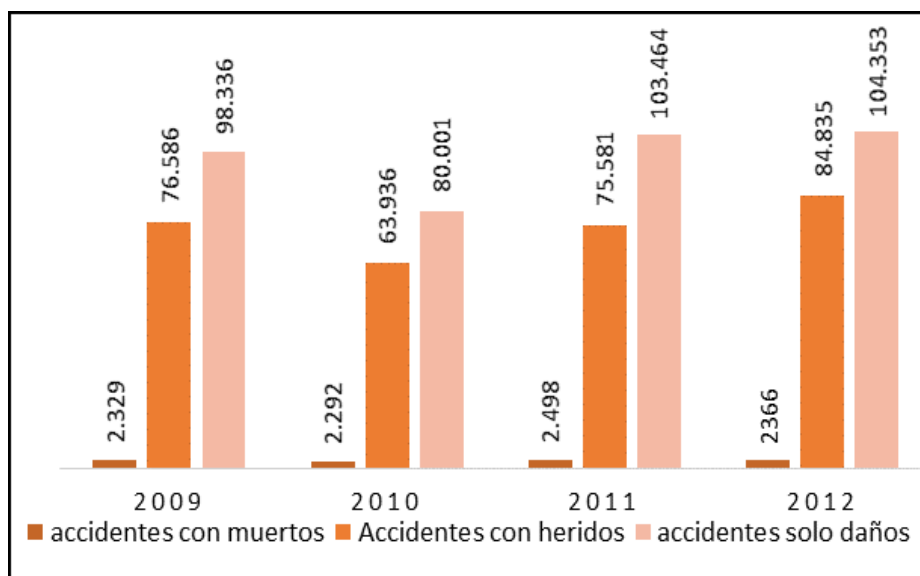
---

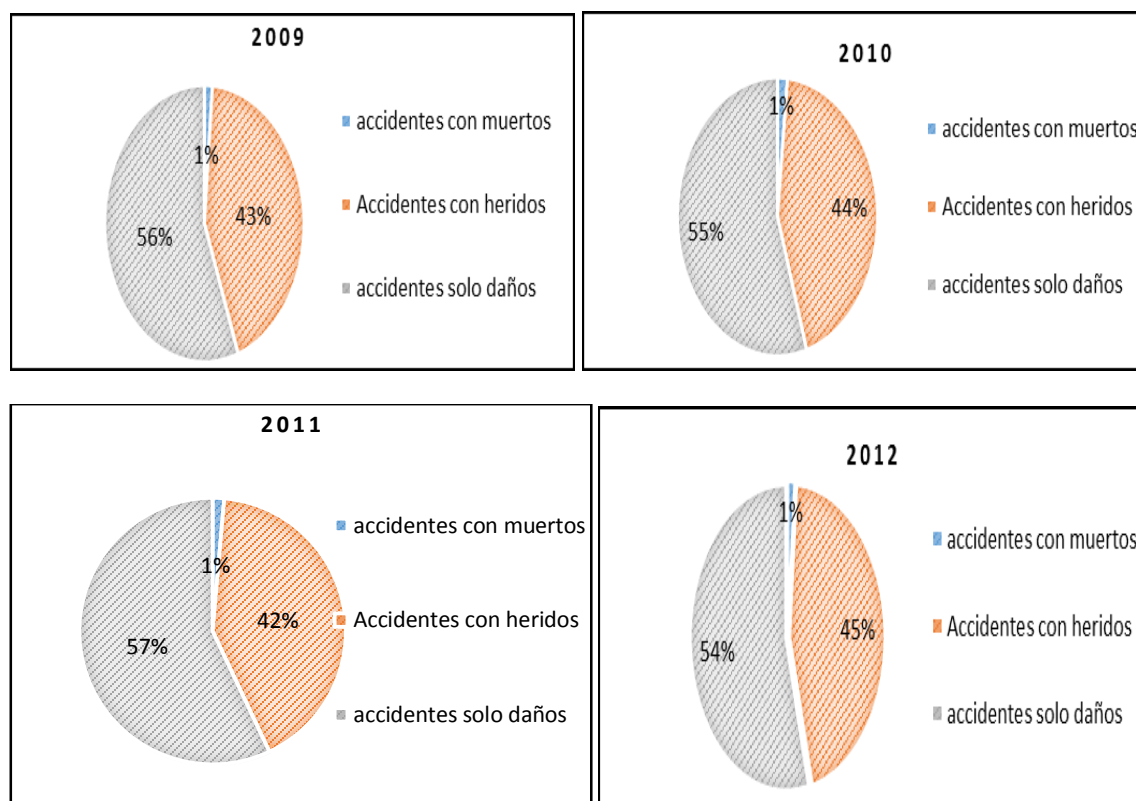
<sup>1</sup> Tomado de *Comportamiento de muertes y lesiones por accidente de transporte, Colombia, 2013 medicina legal*



**Figura 3 Lesiones por accidentes de transporte. Colombia, 2004-2013**  
**Fuente:** Elaboración Propia a partir de datos de Instituto Nacional de Medicina Legal y Ciencias Forenses.

A pesar del aumento presentado en la cantidad de accidentes entre 2011 y 2012, la distribución de los eventos según la gravedad de los accidentes para estos dos años es prácticamente igual. Durante los últimos 4 años, el 44,6% de los accidentes graves ocurridos en el país produjeron lesiones tanto fatales (1,4%) como no fatales (43,2 %). El 55,4% restante corresponde accidentes de solo daños, tal como se muestra en la Figura 4.





**Figura 4 Gravedad de los accidentes – total nacional 2009 – 2012.**  
**Fuente:** Guía técnica para el diseño de las zonas laterales para vías más seguras.

Los reportes de accidentes indican que el 54,1% de los eventos con muertos sucede en zonas urbanas, mientras que el 45,9% en zonas interurbanas. Además, el 7,9% de los accidentes con heridos y el 6,9% de los accidentes con solo daños se produjo en las carreteras del país (Tabla 1). Dado el menor número de accidentes en zonas urbanas por año, los datos sugieren que la gravedad de los accidentes en carreteras suele ser muy superior a la que ocurre en vías urbanas.

Año	Área	Total de accidentes	Accidentes con muertos		Accidentes con heridos		Accidentes solo daños	
			Cantidad	%	Cantidad	%	Cantidad	%
2009	Urbana	163.191	1.224	52,6	70.170	91,6	91.797	93,4
	Carreteras	14.060	1.105	47,4	6.416	8,4	6.539	6,6
	Totales	177.251	2.329	100	76.586	100	98.336	100
2010	Urbana	136.210	1.417	61,8	59.106	92,4	75.687	94,6

Año	Área	Total de accidentes	Accidentes con muertos		Accidentes con heridos		Accidentes solo daños	
			Cantidad	%	Cantidad	%	Cantidad	%
	Carreteras	10.019	875	38,2	4.830	7,6	4.314	5,4
	Totales	146.229	2.292	100	63.936	100	80.001	100
2011	Urbana	146.229	1.373	55,2	69.802	92,4	96.624	93,4
	Carreteras	167.779	1.116	44,8	5.779	7,6	6.840	6,6
	Totales	314.008	2.489	100	75.581	100	103.464	100
2012	Urbana	191.534	1.115	47,1	78.088	92	95.250	91,3
	Carreteras	160.413	1.251	52,9	6.747	8	9.103	8,7
	Totales	351.947	2.366	100	84.835	100	104.353	100
Promedio	Urbana	160.413	1.282	54,1	69.292	92,1	89.840	93,1
	Carreteras	13.729	1.087	45,9	5.943	7,9	6.669	6,9
	Totales	174.142	2.369	100	75.235	100	96.509	100

**Tabla 1 Accidentes por gravedad, según el área de ocurrencia – total nacional 2009- 2012.**

**Fuente:** Guía técnica para el diseño de las zonas laterales para vías más seguras

## 2.3 Objetivo general

La investigación tiene por objeto proponer y estimar modelos econométricos para evaluar la severidad de los accidentes en vías interurbanas colombianas, en función de las características de la infraestructura, del tránsito.

Para lograr este objetivo se propuso una estrategia metodológica para obtener la información necesaria que permitiera estimar los modelos econométricos propuestos; además, especificar y estimar modelos discretos probabilísticos para evaluar la severidad de los accidentes.

## 2.4 Hipótesis

La hipótesis básica que se fundamenta la investigación, consiste en suponer que es posible no sólo formular, sino estimar, modelos econométricos de elección discreta que expliquen la severidad de los accidentes, con consistencia estadística.

## **2.5 Contribuciones**

La investigación aporta al entendimiento del fenómeno de la accidentalidad en las vías interurbanas de Colombia mediante la estimación de modelos desde el punto de vista econométrico, que permitan explicar la influencia de variables como las características del conductor, la vía y el tránsito sobre la severidad de los accidentes. Esta información es esencial para establecer prioridades referentes a políticas e intervenciones para disminuir la accidentalidad y salvar vidas.

## **2.6 Estructura de la tesis**

El documento está estructurado como sigue. El capítulo 2: contextualiza los antecedentes necesarios para obtener un marco de referencia sobre la temática de la investigación y una reseña de los métodos y modelos estadísticos utilizados. El capítulo 3: trata la descripción de la metodología del estudio, incluyendo una breve descripción de la base de datos utilizada. El capítulo 4 expone la conceptualización de los modelos de elección estimados, con sus respectivos resultados. Al final el documento presenta las conclusiones y extensiones que se determinan en la investigación.

### 3 ANTECEDENTES

Son diversas las variables que influyen en los accidentes de tránsito. Entre éstas se destaca la velocidad, la magnitud de los flujos, la capacidad vial, las distancias de visibilidad, las condiciones del pavimento, el diseño geométrico, y el comportamiento de los usuarios. El análisis de la accidentalidad exige conocer la influencia de las variables, realizando un análisis integral de los puntos más críticos, a fin de proponer políticas o intervenciones físicas para disminuir la ocurrencia y severidad de los eventos.

En la Figura 5 se presentan 4 factores muy importantes a tener en cuenta para poder controlar los accidentes de tránsito, y así obtener una disminución en la severidad de los accidentes de tránsito. Estos factores son el ser humano, el vehículo, el ambiente y el entorno (la vía).

#### **Humano:**

El factor humano es una de las variables importantes en los accidentes de tránsito, en la cual pueden interviniendo algunas de las siguientes causas mencionadas:

- Conducir bajo los efectos del alcohol, medicinas y estupefacientes.
- Conducir con fatiga, cansancio o con sueño.
- Realizar maniobras imprudentes y de omisión por parte del conductor, por ejemplo; no respetar los señalamientos viales.
- Conducir a exceso de velocidad (produciendo vuelcos, salida del automóvil de la carretera, derrapes).
- Salud física del conductor (ceguera, daltonismo, sordera).

#### **Vehículo:**

El factor Vehículo también es una de las variables importantes en los accidentes de tránsito, debido a vehículo en condiciones no adecuadas para su operación (sistemas averiados de frenos, eléctricos, dirección o suspensión).



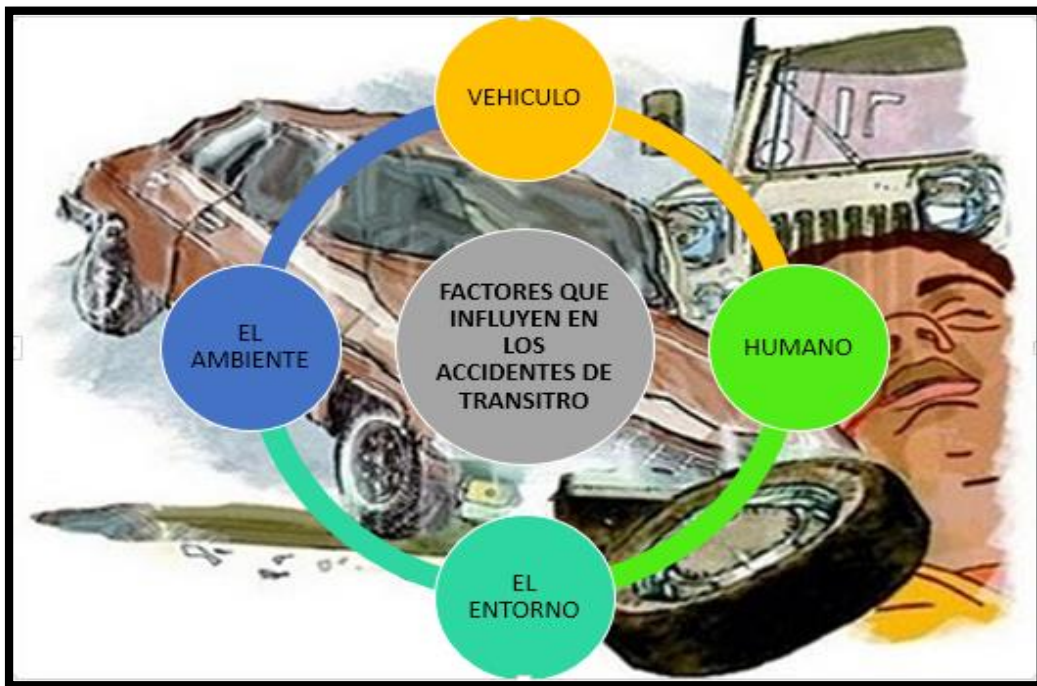
### **Ambiente (Climatológico):**

El factor climatológico también es importante en la ocurrencia de un accidente de tránsito ya que diferentes características del clima pueden causar que el usuario no tenga las mejores condiciones para conducir, debido a que se puede presentar, Niebla, derrumbes, zonas inestables

### **El entorno (La vía):**

Las características del entorno, es una de las variables más importante en la gravedad de un accidente, debido a que no es una variable que pueda corregir el usuario de la vía. Ya que pueden intervenir algunas de las siguientes causas mencionadas:

- Errores de señalamientos viales.
- Carreteras en mal estado o sin mantenimiento (baches, hoyos, pavimento deteriorado).
- Deficientes trazados geométricos, curvas muy cerradas e inseguras.
- La falta de pintura y reflejantes en las líneas centrales y laterales de la carretera.

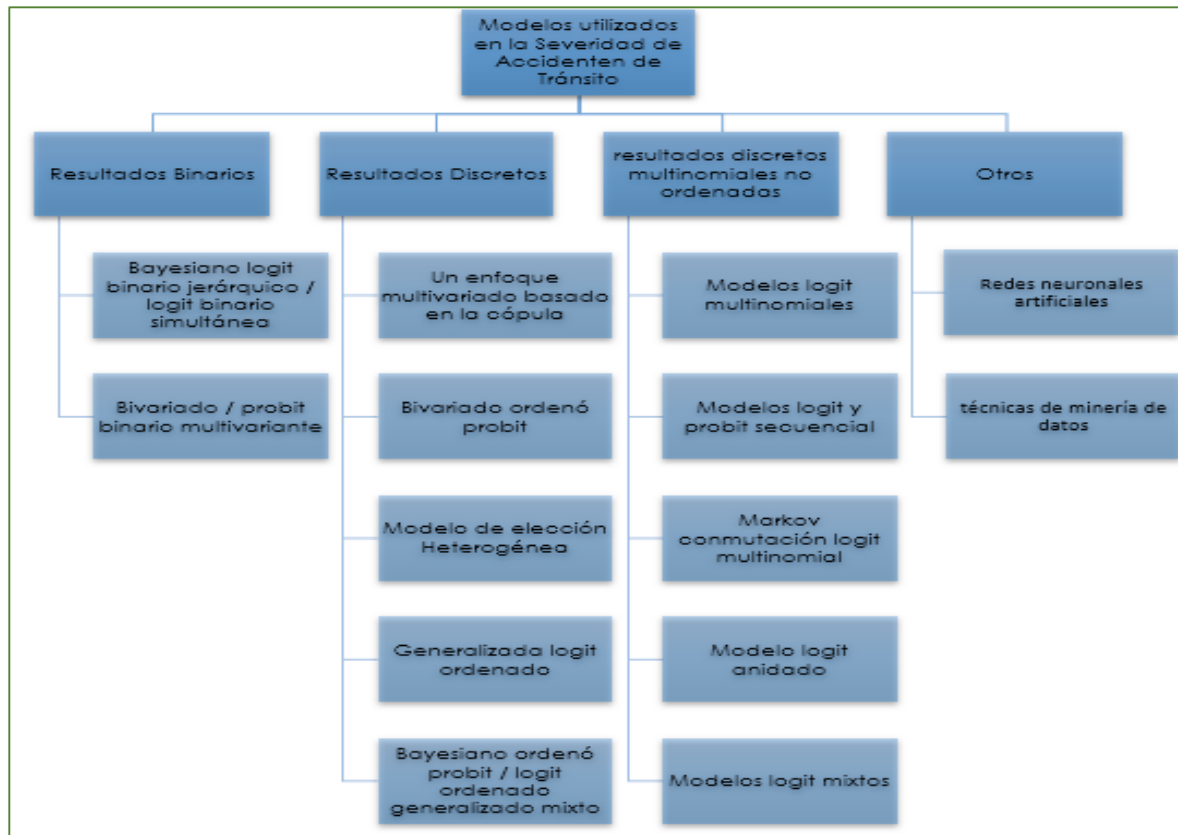


**Figura 5. Factores que influyen en la severidad de los accidentes de tránsito.**  
Fuente: Elaboración propia

Para analizar los datos de severidad en los accidentes se ha aplicado una amplia variedad de técnicas y metodologías estadísticas. La Figura 6 muestra la variedad de técnicas estadísticas que se han utilizado para estudiar la severidad de los accidentes. Los modelos de respuesta discreta logit y probit son utilizados generalmente para explorar la relación entre la gravedad del accidente y sus factores contribuyentes. Un estudio realizado por Savolainen et al. (2011) mostró que los modelos para determinar la gravedad de los accidentes se pueden clasificar como nominales u ordinales. Entre los nominales se encuentran los modelos de respuesta discreta multinomiales no ordenados, siendo los tres más comunes los modelos logit multinomial, logit anidado y logit mixto.

Por otra parte, la gravedad de los accidentes también se caracteriza como una variable ordenada (muertos, heridos o solo daño), usando modelos de respuesta ordenada. Los modelos ordinales clasifican las respuestas siguiendo una escala de orden. Los más comunes son los modelos logit ordinal, probit ordinal y logit mixto ordinal. Sin embargo, este enfoque presenta limitaciones pues el impacto de las variables exógenas sobre las alternativas de gravedad del accidente puede variar. Eluru et al. (2008) se han ocupado de la limitación de los modelos de repuesta ordenada al permitir que los impactos de las variables exógenas puedan variar a través de las alternativas. Diversas investigaciones tanto empíricas como metodológicas utilizando modelos de respuesta ordenada han arrojado resultados positivos, en un enfoque centrado particularmente en la víctima del accidente e intentan aislar los factores de riesgo que contribuyen a su gravedad.

Hay otros modelos alternativos que pueden ser utilizados, como se muestra en la Figura 6, entre los que se incluyen técnicas como redes neuronales y minería de datos, pero que no son tan populares o tan utilizados en la práctica.



**Figura 6 Modelos Utilizados en investigaciones anteriores para evaluar severidad de accidentes de tránsito.**  
**Fuente: (Savolainen et al, 2011)**

- **Modelos de respuesta discreta para evaluar severidad en accidentes de tránsito**

En esta sección se muestra modelos de la familia logit para el análisis de la gravedad de los accidentes.

La accidentalidad en carreteras puede entenderse como un sistema altamente complejo que relaciona factores relacionados con el usuario, el vehículo, la vía y el entorno, pudiendo ocurrir que sólo uno o la interacción entre ellos sea la causa de un evento. Los modelos discretos de la familia logit, pueden emplearse para explorar la relación entre la severidad del accidente y sus factores contribuyentes, tal como lo muestran varias investigaciones. Shankar y Mannering (1996) estudiaron la gravedad de los accidentes entre vehículo/motocicleta utilizando modelos logit multinomial y concluyeron que el uso del casco puede reducir la gravedad de la lesión en accidentes de motociclista. Ulfarsson y Mannering (2004) analizaron la diferencia de la gravedad de la lesiones entre hombres y mujeres concluyendo que golpear una barrera de protección puede causar un nivel de gravedad diferente para ambos géneros. A su vez, Malyskhina y Mannering

(2008) investigaron la influencia del aumento de los límites de velocidad en la gravedad de los accidentes en carreteras del estado de Indiana, Estados Unidos, y concluyeron que mayores límites de velocidad tiene un impacto significativo en la severidad.

Se destaca la investigación de Gkitza y Mannering (2008), quienes utilizaron datos obtenidos de una encuesta aplicada a conductores en India para estimar modelos logit mixto sobre el uso del cinturón de seguridad en vehículos con uno o varios ocupantes. El estudio verificó la hipótesis según la cual el enfoque logit-mixto ofrece una mayor flexibilidad durante la captura de la heterogeneidad-individuo. Concluyen que el comportamiento del conductor se determina según el tipo de vehículo y las características de la carretera donde se encuentre.

La investigación de Milton et al. (2008) utilizó modelos logit mixto para estudiar la gravedad de los accidentes en segmentos de carretera en el estado de Washington, Estados Unidos. Como resultado determinaron que las variables relacionadas con el flujo vehicular como el tránsito promedio diario, la composición vehicular, las características del clima, entre otras, se modelan mejor con modelos de parámetros aleatorios, mientras que las características de la vía como: número de curvas horizontales, la fricción del pavimento se modelan mejor como parámetros fijos.

Malyskhina y Mannering (2010) utilizaron un modelo logit mixto para evaluar la gravedad de accidentes en zonas urbanas y concluyeron que en ellas hay menos probabilidad de accidentes con lesiones, mientras que en zonas interurbanas donde hay altos límites de velocidad, se tiene más probabilidad de que haya accidentes con lesiones más graves.

Yamamoto, Hashiji y Shankar (2008) concluyeron que el sub-registro en los datos no afecta significativamente la estimación de los modelos no ordenados como el logit mixto y el logit multinomial. La razón de esto es que son estructuralmente flexibles. Sin embargo, esto no ocurre con los modelos ordenados, donde la flexibilidad estructural es relevante, ya que permite que diferentes conjuntos de variables independientes puedan ser seleccionadas como predictores significativos de los diferentes niveles de gravedad de la lesión.

Un modelo mixto ordenado para determinar las variables significativas que influyen en los resultados de severidad de lesiones de peatones y ciclistas, fue desarrollado por Eluru et al (2008). Ellos concluyen que peatones y ciclistas de mayor edad, en conjunto con el límite de

velocidad de la carretera, la ubicación geométrica y la hora día, influyen significativamente en el resultado de la severidad de las lesiones en accidentes. También hallaron que los accidentes que ocurren en intersecciones semaforizadas son típicamente menos severos que los que ocurren en otros lugares.

El trabajo de Haque et al (2010) está cimentado sobre el uso de modelos logit binarios para determinar si los motociclistas involucrados en accidentes tienen responsabilidad en la ocurrencia. Encuentran los autores que los factores que contribuyen a la participación en accidentes de motociclistas tienen diferentes efectos dependiendo del lugar del accidente.

Yang et al. (2011) estimaron un modelo probit ordenado para la evaluación de los factores de riesgo asociados con la severidad del accidente en áreas divergentes. Sus resultados indican que los factores más influyentes en la severidad de un accidente son la diferencia de límites de velocidad, el número de carriles, el flujo de tráfico, las condiciones climáticas, el uso de suelo en el entorno de la vía, la presencia de alcohol/drogas, la condiciones de la superficie de la carretera, el número de vehículos implicados y el tipo de colisión.

Kim et al. (2011) y Ye y Lord (2011) desarrollaron un modelo logit mixto para evaluar la gravedad de las lesiones del conductor de un vehículo. Al considerar la influencia de la edad del conductor en los accidentes de autos particulares, concluyen que usuarios de edades avanzadas, mayores a 65 años, tienen mayor probabilidad de sufrir una lesión más grave. También determinaron que otros factores tienen alta incidencia, destacando entre ellos la conducción en estado de alcoholización, el estado de las farolas y las condiciones de iluminación.

### **3.1 Modelo Logit Multinomial MNL**

Estudios anteriores han utilizado el MNL para modelar la severidad de los accidentes de tránsito (Shankar et al., 1996; Carson y Mannering, 2001; Ulfarsson y Mannering, 2004; Kim et al., 2007; Rifaat et al, 2011). Suponiendo que  $P_{jq}$  se refiere a la probabilidad de que un accidente  $q$  de tenga severidad  $j$ , y que se puede representar por:

$$P_{jq} = (U_{jq} \geq U_{jq'}), \forall q' \in Q, q' \neq q \quad (1)$$

Donde  $Q$  denota el conjunto de todos eventos posibles. Se asume que los niveles de severidad se excluyen mutuamente. En este contexto, la función  $U_{jq}$  se puede definir como:

$$U_{jq} = \beta X_{jq} + \varepsilon_{jq} \quad (2)$$

Donde  $\beta$  denota un vector de parámetros estimables,  $X_{jq}$  denota un vector de las características observables que determina la gravedad  $j$  del evento  $q$ .  $\varepsilon_{jq}$  denota un error aleatorio no observable asociado a la severidad  $j$  del evento  $q$ . McFadden (1981) sugirió que si se asume que  $\varepsilon_{jq}$  distribuye Gumbel, se deriva el modelo logit multinomial de la siguiente manera:

$$P_{jq} = \frac{\exp(\delta V_{jq})}{\sum_{A_i \in A_q} \exp(\delta V_{iq})} \quad (3)$$

### 3.2 Modelos de respuesta ordinal

Este tipo de modelos parte de la hipótesis de que los valores de una variable cualquiera pueden clasificarse en un conjunto de categorías ordenadas. Por ejemplo, en el caso de la severidad de un accidente se presentan tres categorías: 1. severidad baja, 2. severidad media, 3. severidad alta.

Nivel de severidad baja (Solo Daños) = calificación 1

Nivel de severidad media (Heridos) = calificación 2

Nivel de severidad alta (Muertos) = calificación 3

Los modelos ordinales estiman la probabilidad de que un evento (en este caso un accidente) se encuentre dentro de unos umbrales arbitrarios que definen un nivel de la variable respuesta (en este caso la severidad).

El modelo es construido alrededor de una regresión latente en el cual se expresa la variable no observada  $y_{jq}^*$  según la ecuación (4). Lo que normalmente se observa es la variable  $y_{jq}$ , que se refiere a la respuesta  $j$  (gravedad del evento) observada en el lugar  $q$ .

$$y_{jq}^* = \theta x_q + \varepsilon_{jq} \quad (4)$$

donde:

$x_q$  = es un vector de variables que determinan el orden discreto de cada evento en el lugar  $q$ .

$\theta$  = es un vector de parámetros estimables,

$\varepsilon_{jq}$  = es un término de error asociado a la respuesta  $j$  y al lugar  $q$ , que en el caso del modelo Logit Ordinal (LO) sigue una distribución logística.

Las variables explicativas son definidas según las características de la vía, del tránsito, del entorno, entre otros. Como se ha indicado, la variable  $y_{jq}^*$  no es observable, pero si podemos saber a qué categorías pertenece dependiendo de los umbrales que se hayan estimado para cada nivel de severidad  $y_{jq}$  Figura 7. Específicamente, si se han definido tres categorías ordenadas para la respuesta de la gravedad de los accidentes  $y_{jq}$ , es posible relacionarlas con las variables latentes  $y_{jq}^*$  de la siguiente manera

$$y_{jq} = \left\{ \begin{array}{ll} 1. \text{Solo Daño} & \text{si } -\infty < y_{jq}^* \leq \tau_1 \\ 2. \text{Herido} & \text{si } \tau_1 < y_{jq}^* \leq \tau_2 \\ 3. \text{Muerto} & \text{si } \tau_2 < y_{jq}^* \leq \infty \end{array} \right\}$$

Donde  $\tau$  es un conjunto de umbrales que define los límites correspondientes de las categorías, y que deben ser estimados.

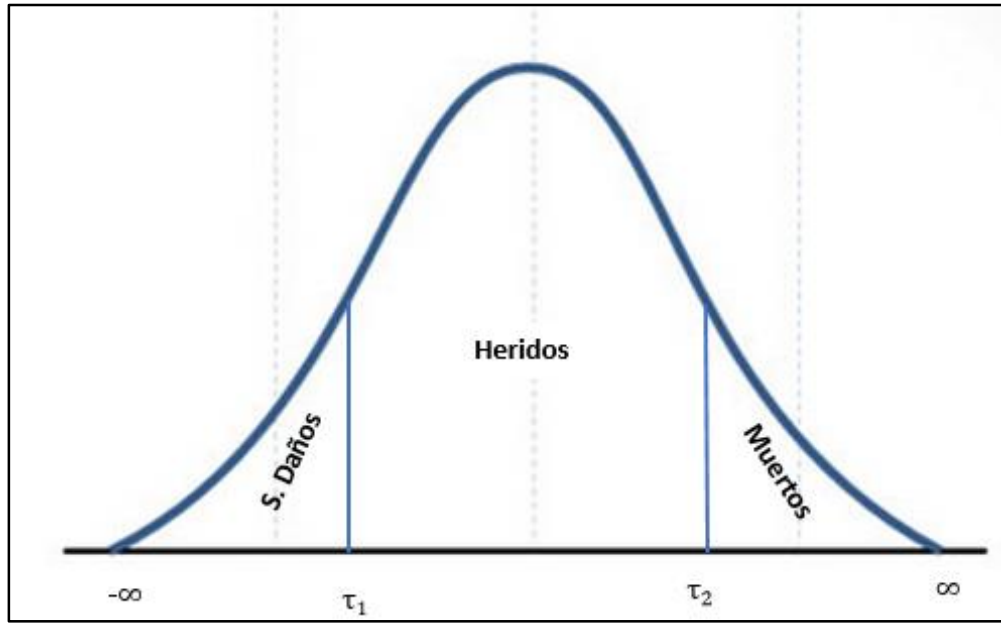


Figura 7 Esquema del modelo ordinal.

Fuente. Elaboración propia

Si  $F$  es la función de distribución acumulada, la probabilidad de observar  $y_{jq}$ , dentro de un indicador discreto o categoría  $c$ , puede escribirse como en la ecuación (5):

$$Pr\{y_{jq} = c | y_{jq}^*\} = F(\tau_c - \theta x_q) - F(\tau_{(c-1)} - \theta x_q) \quad (5)$$

### 3.3 Modelo Logit Mixto Con Componentes De Error.

Un modelo logit mixto, a través de la inclusión de componentes de error, se puede utilizar para capturar correlaciones entre las utilidades de diferentes alternativas. Este modelo puede usarse para capturar una posible correlación entre accidentes con muertos y heridos, los cuales representan mayores consecuencias que un accidente con solo daños. En estos casos las utilidades se especifican como:

$$U_{jq} = \beta X_{jq} + \mu_j z_{jq} + \varepsilon_{jq} \quad (6)$$

Donde  $X_{jq}$  y  $z_{jq}$  son vectores que contienen variables observadas en relación a la alternativa  $j$ ,  $\beta$  es un vector de coeficientes fijos,  $\mu$  es un vector de términos aleatorios con media cero y  $\varepsilon_{jq}$  es de tipo valor extremo iid. Los términos en  $z_{jq}$  son componentes de error que, junto con  $\varepsilon_{jq}$ , definen



la parte estocástica de la utilidad. Es decir, la parte no observada (aleatoria) de utilidad es  $U_{jq} = \mu_j z_{jq} + \varepsilon_{jq}$ , que puede estar correlacionada entre alternativas en función de la especificación de  $z_{jq}$ . Para el modelo logit estándar,  $z_{jq}$  es idénticamente igual a cero, de modo que no hay correlación en la utilidad entre alternativas. Esta falta de correlación da lugar a la propiedad IIA y sus patrones de sustitución restrictivos.

## **4 DESCRIPCIÓN DE LOS DATOS**

En la presente investigación se utilizaron dos bases de datos de naturaleza diferente. La primera comprende los accidentes interurbanos en Colombia entre los periodos 2007 y 2009, indicando su severidad (solo daños, heridos y muertos). La segunda base de datos, corresponde a la descripción de los muertos y heridos por accidentes de tránsito en el periodo 2011- 2012. La diferencia entre las dos base de datos radica en que la primera base se enfoca en describir la severidad de los eventos, mientras que la segunda enfatiza en la caracterización de las víctimas. En otras palabras, en la primera base se puede tener un registro (evento) con una serie de muertos o heridos, mientras que en la segunda se pueden tener varios registros (que corresponden a varios muertos o heridos) para un mismo evento.

Las bases de datos se encuentran en formato Excel. Cada registro fue georreferenciado para poder vincularlo con otra base de datos relacionada con las condiciones geométricas y de tránsito de la vía donde ocurrió el evento. La base de datos que contiene las condiciones geométricas fue proporcionada por el Instituto Nacional de Vías, y contiene características geométricas de las condiciones actuales de un 40% de las vías nacionales del país. La información de esta última base de datos fue obtenida mediante un vehículo equipado con GPS y cámaras que recorrió 11.500 km de la red vial nacional y logró obtener un inventario vial muy detallado. Los datos presentan características de la vía y del entorno como: geometría, uso de suelo, presencia de obstáculos, visibilidad, presencia de intersecciones, entre otros (ver Figura 8).



**Figura 8. Ilustración de los 11.500 km de vía colombiana a las cuales se les realizó el inventario vial**

Los datos de accidentalidad utilizados en este estudio corresponden a los dos periodos que se mencionaron anteriormente, entre 2007-2009 y 2011 -2012. Esta información proviene de los registros del Instituto Nacional de Medicina Legal y Ciencias Forenses, y la Corporación Fondo de Prevención Vial de Colombia (CFPV). En los datos no ha sido incluida la información del 2010 por no encontrarse georrefenciada.

Luego de depurar la información de la base de datos del (2007- 2009) se cuenta con un registro de 11271 accidentes, Cada registro contiene información como:

- Accidente: fecha, departamento, posible causa, tipo de accidente.
- Condición de calzada, tipo de calzada, estado de la calzada
- Especificación de área Urbano/ Rural
- Información acerca del tipo de severidad del accidente.
- Vehículo: tipo de vehículo.

Luego de depurar la segunda base de datos (2011 – 2012) se cuenta con un registro de 1.226 accidentes para la modelación. Cada registro dentro de la base de datos se refiere a la víctima del accidente, si fue muerto o herido, y se cuenta con información de la persona involucrada en el accidente. La anterior información no se tiene para la primera base de datos. Específicamente, cada registro contiene información como:

- Accidente: fecha, departamento, posible causa, tipo de accidente

- Condición de calzada, tipo de calzada, estado de la calzada
- Especificación de área Urbano/ Rural
- Información que indica el tipo de severidad causado por el accidente.
- Característica de la víctima: sexo, calidad (peatón, conductor, pasajero), edad.
- Vehículo: tipo de vehículo.

La información de la magnitud y composición del tránsito y las velocidades proviene de los aforos oficiales del Instituto Nacional de Vías de Colombia y de mediciones del tránsito (flujo y composición vehicular, tipo de vehículo involucrado en los accidentes) efectuadas por la Corporación Fondo de Prevención Vial.

Utilizando un sistema de información geográfico se construyeron las dos bases de datos teniendo en cuenta toda la información disponible en las bases de accidentes, condiciones de la vía y condiciones del tránsito. Así, para cada uno de los registros se tiene información de la gravedad (en tres niveles), del vehículo involucrado, de la persona (en el caso de la base 2011-2012), de la geometría vial, el entorno (uso de suelo) y del tránsito.

#### 4.1. Descripción estadística base de datos 2007 -2009

En esta base de datos, los accidentes se distribuyen así: 6.693 (59,4%) son solo daños, 3.373 (29,9%) son accidentes con heridos y 1.205 (10,7%) son accidentes con muertes. La Tabla 2 muestra las características generales de los datos, mientras que la **¡Error! No se encuentra el origen de la referencia.** muestra las estadísticas de las variables numéricas utilizadas para la estimación de los modelos.

En la Tabla 2 podemos observar las características de los accidentes de la base de datos correspondiente a los años 2007 a 2009 utilizada para la modelación. De ella podemos deducir que la mayoría de los accidentes ocurrieron en vías con anchos de carril mayores a 2,75 m de ancho y donde no había presencia de berma.

<u>Variables</u>		
<u>Características del accidente</u>	<u>Frecuencia</u>	<u>Porcentaje</u>
<u>Severidad</u>		
Solo daño	6693	59.4
Heridos	3373	29.9
Muertos	1205	10.7
<i>Clase de accidentes</i>		
Choque	9247	82.0

<b>Variables</b>		
<b>Características del accidente</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
Atropello	659	5.8
Volcamiento	1270	11.3
Salida de ocupante del vehículo	52	0.5
Incendio	43	0.4
<b>Objeto colisionado</b>		
Vehículo	8456	75.0
Peatón	667	5.9
Semoviente	333	3.0
Objeto fijo	1815	16.1
<b>Condición del pavimento</b>		
Mala	10170	90.2
Bueno	1101	9.8
<b>Anchos de carril</b>		
( $\geq 0$ m a $< 2.75$ m)	199	1.8
( $\geq 2.75$ m a $< 3.25$ m)	1338	11.9
( $\geq 3.25$ m)	9734	86.4
<b>Presencia de berma</b>		
si hay	8749	77.6
no hay	2522	22.4
<b>Existencia de curva</b>		
no hay curva	6784	60.2
si hay curva	4487	39.8
<b>presencia de alumbrado publico</b>		
si hay luz artificial	2149	19.1
no hay luz artificial	9122	80.9
<b>presencia de delineación</b>		
Mala	2657	23.6
Adecuada	8614	76.4

Tabla 2 Características de los datos

Fuente: Elaboración propia.

La **¡Error! No se encuentra el origen de la referencia.** muestra las estadísticas de las variables numéricas utilizadas para la estimación de los modelos.

<b>VARIABLES</b>	<b>UNIDADES</b>	<b>DESCRIPCIÓN</b>	<b>VALORES</b>
Velocidad de Operación	km/h	Máximo	115
		Mínimo	30
		Media	64
		S.D	20
Velocidad de Diseño	km/h	Máximo	95
		Mínimo	30
		Media	50
		S.D	20
TPD	Veh/Dia	Máximo	62.815
		Mínimo	1.137
		Media	12.060
		S.D	9.739

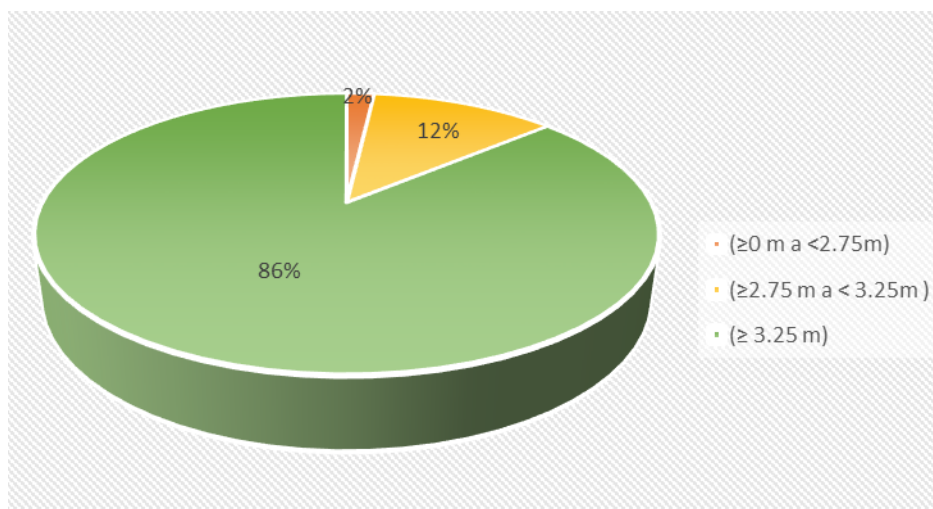
<i><b>VARIABLES</b></i>	<i><b>UNIDADES</b></i>	<i><b>DESCRIPCIÓN</b></i>	<i><b>VALORES</b></i>
Flujo de Motocicletas	Motocicletas/día	Máximo	3.103
		Mínimo	45
		Media	588
		S.D	455
Flujo de Bicicletas	Bicicletas en 100 mt de vía	Máximo	25
		Mínimo	0
		Media	8
		S.D	8
Flujo de peatones	peatones en 100 mt de vía	Máximo	35
		Mínimo	45
		Media	8
		S.D	7
Pendiente Media	%	Máximo	12
		Mínimo	2
		Media	3
		S.D	3

**Tabla 3. Estadísticas descriptivas de las variables base de datos 2007 - 2009**  
**Fuente: Elaboración propia**

Como parte de la preparación de los datos, se clasificaron algunas variables considerando su distribución. Un ejemplo es la categorización de los anchos de carril, que fue clasificada en tres niveles: carriles angostos (ancho <2,75m), carriles medianos (ancho entre 2,75 y 3,25m), y carriles amplios (ancho mayor de 3,25 m). Esta clasificación es sugerida por Abdel-Aty, Chen, y Schott (1998).

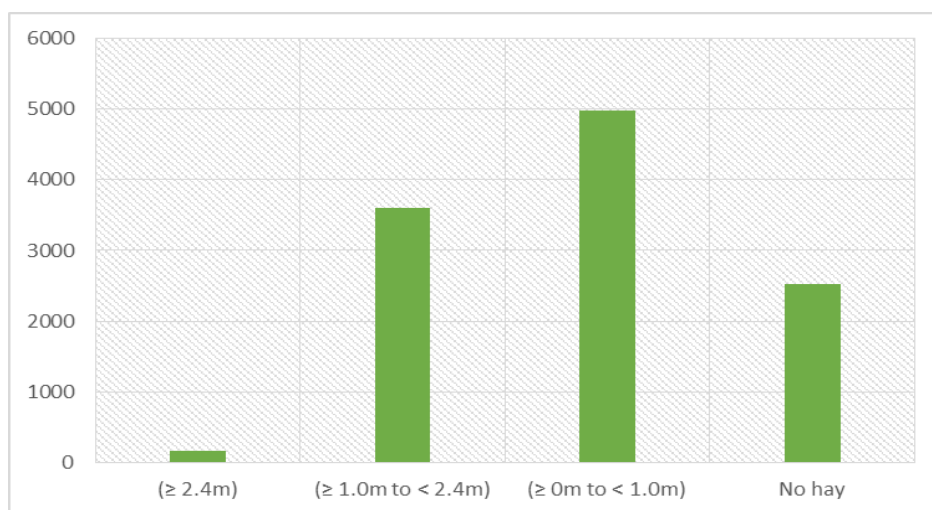
Inicialmente se efectuó un análisis estadístico de los datos para verificar la posible existencia de correlación entre las variables independientes. Las características de los vehículos no han sido incluidas en este estudio debido a que la mayoría de los registros no tenían completa la información que describiera el tipo de vehículo involucrado en el evento.

La **¡Error! No se encuentra el origen de la referencia.** describe otras variables relevantes dentro del análisis. Se observa que la mayoría de los accidentes son choques con otro vehículo (82%). Nótese que la gran cantidad de accidentes ocurren en vías de carriles grandes con un 86,4% (ver Figura 9).



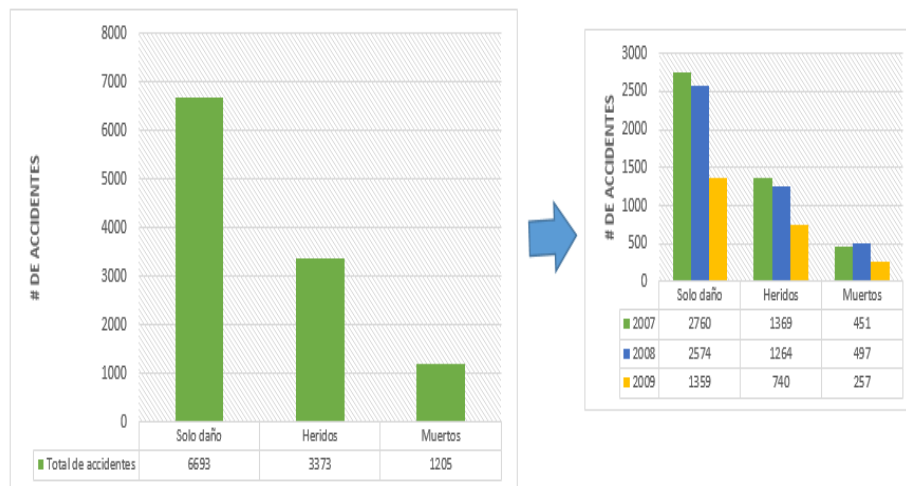
**Figura 9 Distribución porcentual de los anchos de carril de los accidentes.**  
Fuente: Elaboración propia

También se pudo observar que en el 22,4 % de las vías donde suceden accidentes no existe berma (ver Figura 10).



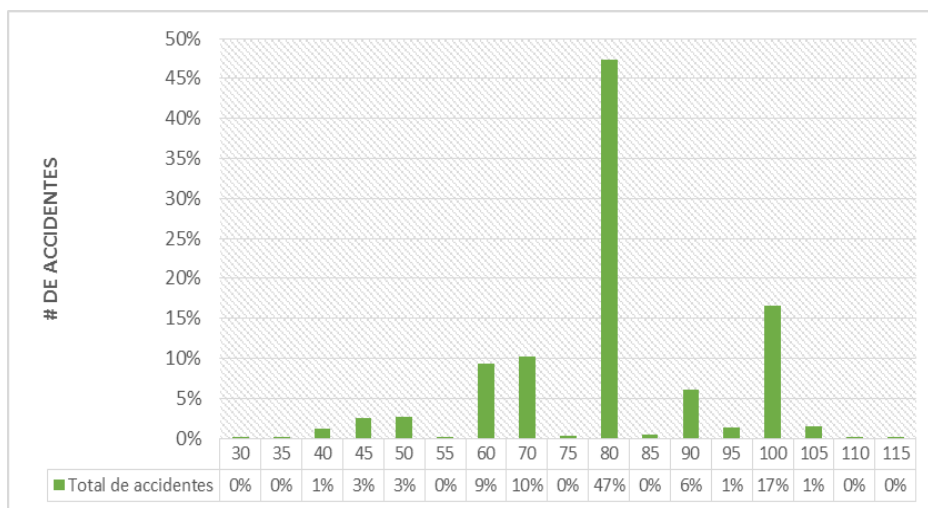
**Figura 10 Distribución porcentual de los anchos de berma en los accidentes.**  
Fuente: Elaboración propia

En la Figura 11 podemos observar que los accidentes con solo daños predominan en el periodo 2007 a 2009, seguido de los heridos y muertos.



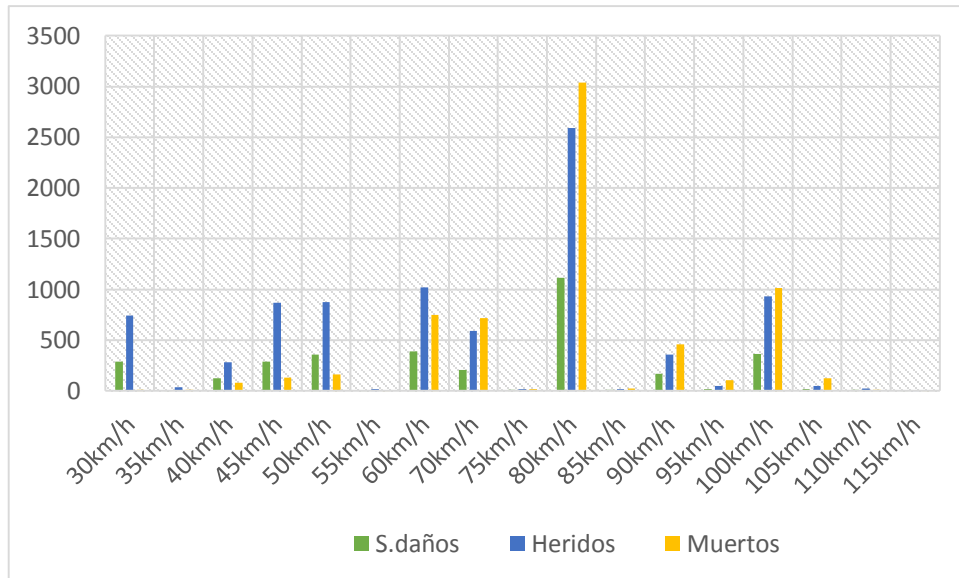
**Figura 11. Distribución porcentual de la gravedad de los accidentes por año.**  
Fuente: Elaboración propia

Es interesante notar que cerca del 55% de los accidentes de tránsito ocurrieron en vías cuyas velocidades de operación se encontraban entre 80 km/h y 100 km/h (ver Figura 12). La mayor parte de estos accidentes terminaron con muertos y heridos.



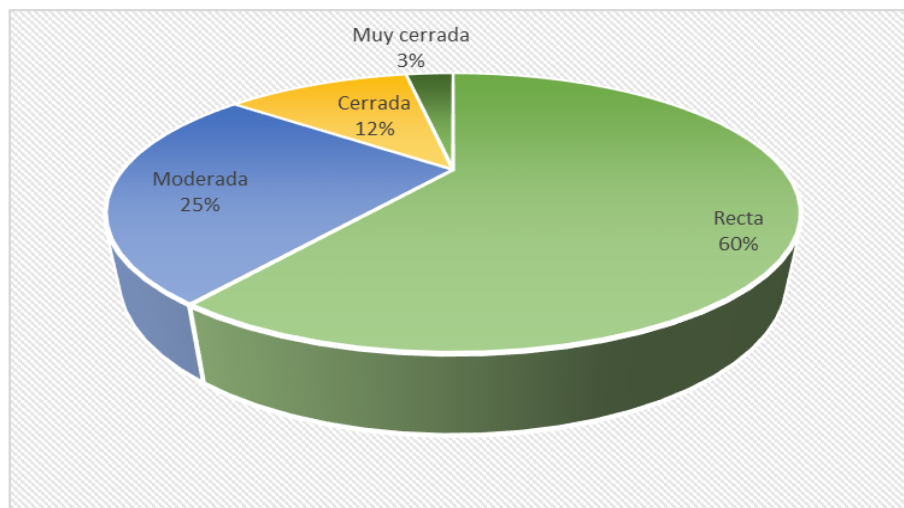
**Figura 12 Distribución de las velocidades de operación de los accidentes.**  
Fuente: elaboración propia, base de datos utilizada para la modelación





**Figura 13 Distribución de las velocidades de operación de los accidentes según la gravedad del accidente.**  
**Fuente: elaboración propia, base de datos utilizada para la modelación**

En la Figura 14 se puede observar que el 60% de los accidentes se presentan en zonas rectas y que un 25% se presentó en vías con curvas moderadas. Lo anterior sugiere que posiblemente los tramos rectos permiten desarrollar más velocidad y las colisiones que se presentan en ellas tienen mayor consecuencia.



**Figura 14 Distribución de presencia de curvas en los accidentes.**  
**Fuente: Elaboración propia**

En la **¡Error! No se encuentra el origen de la referencia.** se hace una descripción de las variables utilizadas en los modelos. Algunas han sido tomadas como variables mudas y otras

como numéricas. La descripción en la tabla también nos indica cuales fueron los criterios para definir dichas variables.

<b>Tipo de Variables</b>	
<b>Condición de la carretera</b>	<b>Descripción</b>
Transito promedio Diario Anual	Vehículos por día/1000
Ancho de carril	1 Si el carril es >3,25mt, 0 si el carril es <3,25mt
Presencia de berma derecha	1 Si no hay berma, 0 si es hay presencia de berma
Existencia de curva	1 Si hay curva, 0 en caso contrario
Velocidad de operación	Varia entre 30 km/h y 110 km/h
Señalización horizontal en buen estado	1 Si hay iluminación noturna, 0 si no hay
Resistencia al deslizamiento	1 si es Adecuada, 0 si no lo es
Presencia de demarcación	1 en buen estado, 0 en caso contrario
Presencia de objeto lateral en la vía	1 si hay presencia de objeto lateral, 0 en otro caso
Tipo de área	1 en zona rural, 0 en otro caso
<b>Información del Accidente</b>	
Severidad	1=Solo daños, 2= Heridos 3= Muertos
Peatón	1 si el afectado es peatón, 0 en otro caso
Motociclista	1 si el afectado es motociclista, 0 en otro caso

**Tabla 4 Características y tipos de las variables a utilizar en los modelos para la base de datos 2011- 2012**

**Fuente: Elaboración propia**

#### **4.2. Descripción estadística base de datos 2011 -2012**

Para la construcción de la base de datos 2011 – 2012 se realizó la misma metodología utilizada para la base anterior. Cabe recordar que cada registro en esta base de datos representa un muerto o un herido. Para la modelación se cuenta con 1226 registros de heridos o muertos en accidentes. Cada registro se refiere a la víctima del accidente.

En esta base de datos, los accidentes se distribuyen así: 693 (56,5%) fueron Heridos, 533 (43,5%) son accidentes con Muertos. En la Tabla 5 **Tabla 5** podemos observar las características de los accidentes de la base de datos correspondiente a los años 2011-2012 utilizados para la modelación.

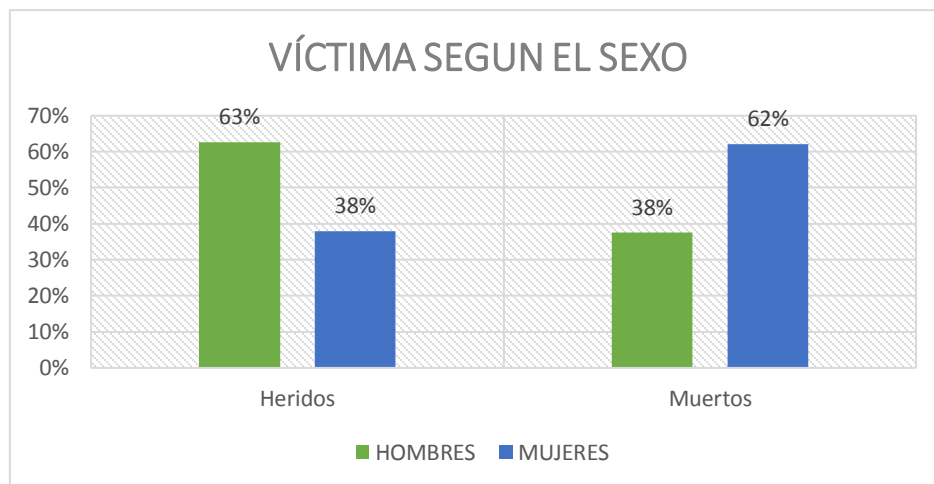
<b>Variables</b>		
<b><u>Características del accidente</u></b>	<b><u>Frecuencia</u></b>	<b><u>Porcentaje</u></b>
<b><u>Severidad</u></b>		
Heridos	693	0.565
Muertos	533	0.435
<b><i>Características de Vehículo</i></b>		
Automóvil	424	0.35
Bus-buseta	128	0.10
Camioneta	42	0.03
Motocicleta	535	0.44

<b>Variables</b>		
<b>Características del accidente</b>	<b>Frecuencia</b>	<b>Porcentaje</b>
<b>Severidad</b>		
Maquinaria tractor	42	0.03
Bicicleta	55	0.04
<b>Características de la víctimas</b>		
Peatón	170	0.14
Conductor	319	0.26
Pasajero	222	0.18
Motociclista	515	0.42
<b>Características del accidente</b>		
Atropello	294	0.24
Caída de ocupante	44	0.04
Choque con objeto fijo	550	0.45
Caída al precipicio	36	0.03
Incendio	3	0.00
Volcamiento	81	0.07
Choque	218	0.18
<b>Sexo de víctima</b>		
Masculino	928	0.76
Femenino	298	0.24
<b>Nivel De Educación de víctima</b>		
Primaria	304	0.25
Secundaria	515	0.42
Técnico	104	0.08
Profesional	303	0.25
<b>Estado Civil de víctima</b>		
Casado	35.00	0.03
Soltero	625.00	0.51
Divorciado	28.00	0.02
Unión libre	524.00	0.43
Viudo	14	0.01
<b>Presencia de berma</b>		
Si hay	976	0.796
No hay	250	0.204
<b>Existencia de curva</b>		
No hay curva	722	0.589
Si hay curva	504	0.411

**Tabla 5 Características de los datos**  
**Fuente: Elaboración propia.**

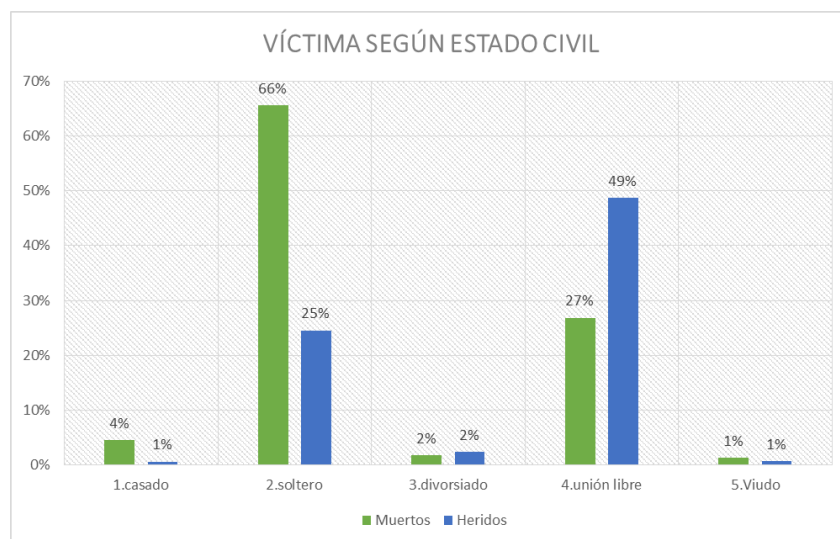
De ella podemos deducir que la mayoría de las víctimas son hombres con un 76% y las mujeres con un 24%. Es de notar que aunque las mujeres presentan un menor porcentaje de las víctimas tienen a tener lesiones más graves (ver Figura 15).

La Tabla 5 describe otras variables relevantes dentro del análisis. Se observa que la mayoría de los accidentes son choques con un objeto con un 45%, atropello con un 24% y seguido se encuentra choque con otro vehículo con un 18%.



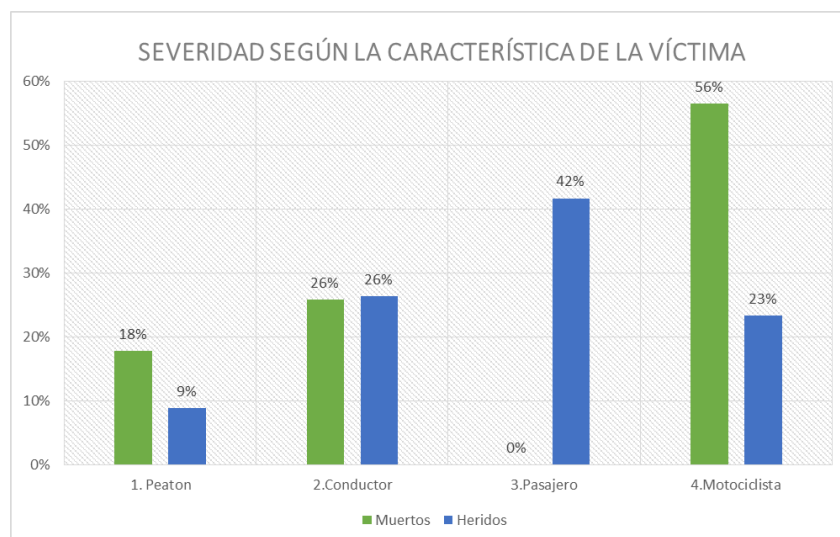
**Figura 15. Distribución de la severidad de accidente según el tipo de sexo.**  
Fuente: Elaboración propia

Es interesante notar que cerca del 66% de los muertos son solteros y un 25% son heridos. Otro porcentaje alto se presenta en los usuarios que viven en unión libre, de los cuales el 49% son heridos y el 27% son muertos (ver Figura 16).



**Figura 16 Distribución De La Severidad De Accidente Según El Estado Civil**  
Fuente: Elaboración propia

En la Figura 17 se puede observar que el 56% y el 18% de las víctimas muertas eran motociclistas y peatones. También se puede observar que los pasajeros de los vehículos solo presentan víctimas con heridos ocupando el mayor porcentaje de heridos con un 42%.



**Figura 17 Distribución De La Severidad De Accidente Según la característica de la víctima**  
**Fuente: Elaboración propia**

En la Tabla 6 se hace una descripción de las variables utilizadas en los modelos. Algunas han sido tomadas como variables mudas y otras como numéricas. La descripción en la tabla también nos indica cuales fueron los criterios para definir dichas variables.

Tipo de Variables	
Condición de la carretera	
Transito promedio Diario Anual	Vehículos por día/1000
Presencia de berma derecha	1 Si no hay berma, 0 si es hay presencia de berma
Existencia de curva	1 Si hay curva, 0 en caso contrario
Información de la víctima	
Sexo	1 Si es masculino ,0 si es femenino
Edad	1 si es mayor o igual 40 años, 0 si es menor a 40 años
Estado civil unión libre	1 Si es soltero, 0 en caso contrario
Estado civil soltero	1 Si vive en unión libre, 0 en caso contrario
Educación básica	1 si solo tiene primaria, 0 en otro caso
Información del Vehículo	
Moto	1 si es Moto, 0 en otro caso
Auto	1 si es Auto, 0 en otro caso
Bus	1 si es Bus, 0 en otro caso
Información del Accidente	
Severidad	1=Heridos, 2= Muertos
Peatón	1 si el afectado es peatón, 0 en otro caso
Conductor	1si el afectado es conductor, 0 en otro caso
Exceso de alcohol	1si el afectado presenta exceso de alcohol, 0 en otro caso

**Tabla 6 Características y tipos de las variables a utilizar en los modelos para la base de datos 2011- 2012**  
**Fuente: Elaboración propia**

Realizando una exploración de los datos con una matriz de correlación se pudo observar que no se encuentra correlación entre las variables de la base de datos 2011 -2012. La matriz de correlación la podemos encontrar en los anexos.

## 5. DESCRIPCIÓN DE LOS MODELOS

En el capítulo se presentan los resultados del proceso de modelación de las variables que inciden en la severidad de los accidentes haciendo usos de los datos obtenidos de las bases anteriormente descritas.

Para llegar a la obtención de los modelos de predicción se realizó un pre-procesamiento de los datos en el cual se eliminaron redundancias e inconsistencias para luego normalizar los datos para facilitar las tareas posteriores.

Es importante anotar que los factores que inciden en la severidad de los accidentes son dinámicos y cambian continuamente. Las características de las vías y del entorno dependen de las intervenciones que se hagan sobre ellos. De igual manera, la magnitud y composición del tránsito varían en el tiempo.

### 5.1. Modelos estimados de la severidad de los eventos con la base de datos (2007 -2009).

Utilizando la base de datos del periodo de (2007 – 2009) se realizaron tres modelos los cuales se presentan a continuación: Modelos Logit Multinomial, Ordinal y Logit Mixtos con Componentes de Error.

- **Modelo Logit Multinomial MNL**

Estudios anteriores han utilizado el MNL para modelar la severidad de los accidentes de tránsito (Shankar et al., 1996; Carson y Mannering, (2001); Ulfarsson y Mannering, (2004); Kimet al, x2007;. Rifaat et al, 2011). Suponiendo que  $P_{jq}$  se refiere a la probabilidad de que un accidente  $q$  de tenga severidad  $j$ , y que se puede representar por:

$$P_{jq} = (U_{jq} \geq U_{jq'}), \forall q' \in Q, q' \neq q \quad (7)$$

Donde  $Q$  denota el conjunto de todos eventos posibles. Se asume que los niveles de severidad se excluyen mutuamente. En este contexto, la función  $U_{jq}$  se puede definir como:

$$U_{jq} = \beta X_{jq} + \varepsilon_{jq} \quad (8)$$

Donde  $\beta$  denota un vector de parámetros estimables,  $X_{jq}$  denota un vector de las características observables que determina la gravedad  $j$  del evento  $q$ .  $\varepsilon_{jq}$  denota un error aleatorio no observable asociado a la severidad  $j$  del evento  $q$ . McFadden (1981) sugirió que si se asume que  $\varepsilon_{jq}$  distribuye Gumbel, se deriva el modelo logit multinomial de la siguiente manera:

$$P_{jq} = \frac{\exp(\delta V_{jq})}{\sum_{A_i \in A_q} \exp(\delta V_{iq})} \quad (9)$$

Donde  $V_{jq}$  es una combinación lineal de parámetros y características observables que determinan la gravedad  $j$  del evento  $q$ . Las variables utilizadas en la especificación de los modelos se presentan en la Tabla 7.

<i>VARIABLE</i>	<i>DEFINICIÓN</i>
<b>VARIABLES EN HERIDOS</b>	
X1 - Ancho de carril mayor 3,25 metros	Anchos de carril incluye la descripción de la tabla 4
X2 = No Hay Presencia de berma	No hay presencia de berma en el lugar del accidente. Incluye la descripción de la tabla 4.
X3 = existencia de curva en 100 m de vía	Existencia de curva en el lugar del accidente.
X4 = TPDA	Flujo promedio diario anual /1000.
X5 =señalización horizontal en buen estado	Es la presencia de señalización horizontal de la vía donde ocurrió el accidente
X6 = porcentaje de motos en la vía	Flujo promedio diario anual /100
X7 = presencia de peatones en 100 metros de vía	Presencia de peatones en 100 metros de vía.
X8 = presencia de bicicletas en 100 metros de vía	Presencia de bicicletas en 100 metros de vía
X9 = pendiente mayor al 4%	pendiente mayor al 4%
X10 = pendiente de intersección en 100 metros de vía	Si hay o no presencia de intersección en 10 metros de vía
<b>VARIABLES EN MUERTOS</b>	
X12 = No Hay Presencia de berma	No hay presencia de berma en el lugar del accidente. Incluye la descripción de la tabla 4.
X13 = existencia de curva en 100 m de vía	Existencia de curva en el lugar del accidente.

<i>VARIABLE</i>	<i>DEFINICIÓN</i>
X14 = TPDA	Flujo promedio diario anual /1000.
X15 = señalización horizontal en buen estado	Es la presencia de señalización horizontal de la vía donde ocurrió el accidente
X16 = porcentaje de motos en la vía	Flujo promedio diario anual /100
X17 = presencia de peatones en 100 metros de vía	Presencia de peatones en 100 metros de vía.
X18 = presencia de bicicletas en 100 metros de vía	presencia de bicicletas en 100 metros de vía
X19 = pendiente mayor al 4%	pendiente mayor al 4%
X20 = pendiente de intersección en 100 metros de vía	Si hay o no presencia de intersección en 10 metros de vía

**Tabla 7 Definición de las variables a utilizar en la modelación**  
Fuente: Elaboración propia

La combinación lineal entre las variables y los coeficientes para describir la gravedad  $j$  de cada evento  $q$  se presenta en las ecuaciones a continuación:

$$V_{Solodaños} = ASC1 * one \quad (10)$$

$$\begin{aligned}
 V_{Heridos} = & ASC2 * one + B2_{berma} * X2_{berma} + B3_{curva} * X3_{curva} + B4_{TPDA} \\
 & * X4_{TPDA} + B5_{SeñalizaciónH} * X5_{SeñalizaciónH} + B6_{flujo demoto} \\
 & * X6_{flujo demoto} + B7_{FlujodePeatones} * X7_{FlujodePeatones} \\
 & + B8_{Flujodebicicletas} * X8_{Flujodebicicletas} + B9_{Pendienteterreno} \\
 & * X9_{Pendienteterreno} + B10_{P.Intersección} * X10_{P.Intersección}
 \end{aligned} \quad (11)$$

$$\begin{aligned}
 V_{Muertos} = & ASC3 * one + B1_{AnchoCarril} * X1_{AnchoCarril} + B12_{Berma} * X12_{Berma} \\
 & + B13_{Curva} * X13_{Curva} + B14_{Flujo} * X14_{Flujo} + B15_{SeñalizaciónH} \\
 & * X15_{SeñalizaciónH} + B16_{FlujodeMoto} * X16_{FlujodeMoto} \\
 & + B17_{FlujodePeatones} * X17_{FlujodePeatones} + B18_{Pendienteterreno} \\
 & * X18_{Pendienteterreno} + B19_{p.intersección} * X19_{P.Intersección}
 \end{aligned} \quad (12)$$

Para cada severidad se definieron constantes específicas. Por temas de identificación del modelo, se fijó en cero la constante específica para la alternativa Solo daños (ASC1). Los coeficientes de las variables explicativas se definieron como específicas a las alternativas de accidentes con



heridos y con muertes para capturar posibles diferencias en la influencia de estas variables sobre los diferentes niveles de severidad.

La Tabla 8 resume los resultados de la estimación de los coeficientes del modelo logit Multinomial.

<i>MNL</i>			
	Coeficiente	E. Estándar	Test t
Constantes específicas			
<b>Solo daños</b>	0	Fijo	
<b>Heridos</b>	-0,836	0,0652	-12,81
<b>Muertos</b>	-2,01	0,129	-15,57
<b>Características de la vía</b>			
<b>Heridos</b>			
No Hay Presencia de berma	0,127	0,0528	2,40
Presencia de curva en 100 metros de vía	-0,176	0,0470	-3,74
TPDA	-0,173	0,0734	-2,36
porcentaje de motos en la vía	0,733	0,233	3,15
señalización horizontal en buen estado	-0,0748	0,0554	-1,35
presencia de peatones en 100 metros de vía	0,183	0,0329	5,55
presencia de bicicletas en 100 metros de vía	0,186	0,0407	4,58
pendiente mayor al 4%	-0,273	0,0488	-5,59
<b>Muertos</b>			
<b>Características de la vía</b>			
Ancho de carril mayor 3.25 metros	0,318	0,109	2,91
No Hay Presencia de berma	0,298	0,0863	3,45
presencia de curva en 100 metros de vía	-0,339	0,0711	-6,11
TPDA	-0,426	0,111	-3,83
señalización horizontal en buen estado	-0,217	0,0878	-2,47
porcentaje de motos en la vía	0,733	0,233	3,15
presencia de peatones en 100 metros de vía	0,107	0,0494	2,16
presencia de bicicletas en 100 metros de vía	0,0479	0,0604	4,7
pendiente mayor al 4%	-0,397	0,0758	-5,23
<b>Parámetros</b>		19	
<b>Tamaño de la Muestra</b>		11271	
<b>Log-verosimilitud</b>		-10.087,7	

**Tabla 8 Modelo Multinomial Estimado**

**Fuente: elaboración propia.**

Cabe destacar que inicialmente se ensayaron otras variables en los modelos. Entre las variables ensayadas, se encuentra la presencia de intersección cada 100 metros de vía, la presencia de iluminación artificial, el tipo de vehículo involucrado en el accidente, pero fueron retiradas del modelo por ser poco significativas. Se encontró además que la velocidad de operación está fuertemente correlacionada con los anchos de carril, por lo cual se decidió incluir solo esta última.

Puede notarse que en el modelo presentado en la Tabla 8 la mayoría de las variables incluidas resultaron relevantes, significativas y con el signo esperado. Los resultados indican que en vías que no presentan bermas hay más probabilidad que en caso de ocurrir un accidente este sea de severidad alta. Lo anterior se explica por el hecho de que el espacio de maniobra del conductor es más pequeño.

La variable ancho de carril solo fue incluida en muertos, porque sólo resultó ser significativa en este tipo de severidad. Tener carriles más anchos aumenta la probabilidad que la severidad del accidente sea grave. Aunque pudiera pensarse que este es un resultado contra-intuitivo, lo que se piensa que puede ocurrir es que en lugares con carriles más anchos los conductores conducen a mayor velocidad haciendo que la severidad de los accidentes sea mayor.

Lo contrario ocurre con la presencia de curvas en la vía, que disminuyen la probabilidad de que el accidente resulte en severidad alta. Es de mencionar que la velocidad no ha sido incluida en el modelo, ya que está correlacionada con muchas de las variables incluidas. Es claro que la velocidad tiene alta significancia en la severidad del accidente, dado que la energía del impacto es mayor y que esta también influencia en los tiempos de reacción.

Otras variables interesantes son el flujo de motociclistas, peatones y bicicletas en la vía, los cuales a medida que van aumentando, mayor es la probabilidad de que la ocurrencia de un accidente termine en severidades más altas. Lo anterior porque el nivel de exposición de los usuarios más vulnerables aumenta. Entretanto puede notarse que hay menor propensión a accidentes graves en las vías de alto flujo ya que potencialmente los usuarios pueden optar por circular a velocidades más bajas o con mayor precaución por la presencia de otros vehículos. .

El modelo indica que la señalización horizontal en buen estado disminuye la probabilidad de accidentes con muertes.

- **Modelo Logit Ordinal**

Los modelos logit ordinal son de gran utilidad para la modelación de la gravedad de accidentes de tránsito. Establecen la especificación de umbrales para definir los límites en los que se encuentra cada categoría de severidad. Se define una variable no observada en función de variables explicativas que es utilizada como base para la modelación ordinal de los datos, y que es específica como sigue:

(13)

$$\begin{aligned}
ALT1 = & B1_{AnchoCarril} * X1_{AnchoCarril} + B12_{Berma} * X12_{Berma} + B3_{curva} * X3_{curva} \\
& + B4_{TPDA} * X4_{TPDA} + B5_{SeñalizaciónH} * X5_{SeñalizaciónH} \\
& + B6_{flujodemoto} * X6_{flujodemoto} + B7_{FlujodePeatones} * X7_{FlujodePeatones} \\
& + B8_{Flujodebicicletas} * X8_{Flujodebicicletas} + B9_{Pendienteterreno} \\
& * X9_{Pendienteterreno} + B10_{P.Intersección} * X10_{P.Intersección}
\end{aligned}$$

Se establecieron tres categorías así:

1. Solo Daño
2. Heridos
3. Muertos

Los resultados de estimación de los parámetros del modelo se presentan en la Tabla 9

<i>Logit Ordinal</i>			
VARIABLES	Coefficiente	E. estándar	Test t
Heridos	0	fijo	
Ancho de carril	0,230	0,155	1,48
Presencia de berma	0,199	0,0476	4.18
existencia de curva	-0,224	0,0417	-5.37
TPDA	-0,256	0,0649	-3,95
señalización horizontal en buen estado	-0,126	0,0497	-2,54
presencia de peatones en 100 metros de vía	0,0138	0,00292	4.74
presencia de bicicletas en 100 metros de vía	0,0114	0,00364	3,12
Presencia de flujo de motocicletas	0,000437	0,00133	3.28
pendiente mayor al 4%	-0,296	0,0437	-6,78
presencia de intersección en la vía	-0,0997	0,0679	-1.48
$\tau_1$	0,598	0,174	3,44
$\tau_2$	2,37	0,175	13,51
Parámetros			12
Observaciones			11271
Log-verosimilitud			-10.115,50

Tabla 9 Modelo Logit Ordinal

Fuente: Elaboración propia.

La Tabla 9; **Error! No se encuentra el origen de la referencia.** presenta los resultados obtenidos del modelo logit ordinal. Se determinó que a medida que los anchos de carriles son mayores a 3.25 m y si no hay presencia de berma a la derecha del conductor, aumenta la probabilidad de

que la gravedad del accidente sea más severo. Lo anterior puede estar correlacionado con la velocidad, ya que en esta vías el conductor puede percibir un nivel de confianza más alto que cuando transita por carriles angostos. Este nivel de confianza puede conllevar a que se presenten velocidades más altas lo que hace que al ocurrir un accidente este sea más severo.

Otra variable que puede estar correlacionada con la velocidad es la existencia de curva cada 100 mt de vía, ya que si hay presencia de curvas los conductores tiende a estar más alerta y a disminuir la velocidad. Lo anterior causa que en caso de ocurrencia de un accidente la severidad tienda a ser menor.

A medida que el flujo vehicular aumenta disminuye la severidad de los accidentes. Esto debe tener cierta relación con la velocidad de los vehículos. A mayor flujo los vehículos tienden a ir más perceptivos y a bajas velocidades, lo que les permite estar pendientes de las condiciones de la vía. Lo contrario sucede en vías donde el flujo vehicular disminuye.

Otra variable interesante es el flujo de motociclistas, peatones y bicicletas en la vía, los cuales a medida que van aumentando los flujos, mayor es la probabilidad de que la ocurrencia de un accidente termine en severidades más altas.

El modelo nos arroja que cuando la vía presenta una buena señalización horizontal la severidad de los accidentes disminuye. También se puede deducir que la severidad de los accidentes disminuye cuando se presentan pendientes mayores al 4% esto puede también puede tener relación con la velocidad, ya que en pendientes más bajas o terrenos planos la velocidad tienden hacer más altas que en terrenos ondulados o montañoso.

Puede notarse que en vías de carriles más anchos, la no presencia de berma, la presencia de peatones, bicicletas y motociclistas, influyen significativamente en la severidad de los accidentes en las carreteras colombianas. En contraste, la existencia de una buena señalización, la presencia de curvas en la vía y los terrenos con pendientes mayores al 4%, influye significativamente en la disminución de la severidad de un accidente.

- **Modelo Logit Mixto Con Componentes De Error.**

Los modelos logit Mixto con componentes de Error se utilizan cuando se trata de analizar patrones de correlación entre las funciones de utilidad de las alternativas utilizadas en la modelación. Las funciones de utilidad se describen a continuación:

$$V_{\text{Solodaños}} = ASC1 * one \quad (24)$$

$$\begin{aligned} V_{\text{Heridos}} = & ASC2 * one + B2_{\text{berma}} * X2_{\text{berma}} + B3_{\text{curva}} * X3_{\text{curva}} + B4_{\text{TPDA}} \\ & * X4_{\text{TPDA}} + B5_{\text{SeñalizaciónH}} * X5_{\text{SeñalizaciónH}} + B6_{\text{flujodemoto}} \\ & * X6_{\text{flujodemoto}} + B7_{\text{FlujodePeatones}} * X7_{\text{FlujodePeatones}} \\ & + B8_{\text{Flujodebicicletas}} * X8_{\text{Flujodebicicletas}} + B9_{\text{Pendienteterreno}} \\ & * X9_{\text{Pendienteterreno}} + \sigma \end{aligned} \quad (15)$$

$$\begin{aligned} V_{\text{Muertos}} = & ASC3 * one + B1_{\text{AnchoCarril}} * X1_{\text{AnchoCarril}} + B12_{\text{Berma}} * X12_{\text{Berma}} \\ & + B13_{\text{Curva}} * X13_{\text{Curva}} + B14_{\text{Flujo}} * X14_{\text{Flujo}} + B15_{\text{SeñalizaciónH}} \\ & * X15_{\text{SeñalizaciónH}} + B16_{\text{FlujodeMoto}} * X16_{\text{FlujodeMoto}} \\ & + B17_{\text{FlujodePeatones}} * X17_{\text{FlujodePeatones}} + B18_{\text{Pendienteterreno}} \\ & * X18_{\text{Pendienteterreno}} + \sigma \end{aligned} \quad (16)$$

<i>Mixto Con Componentes De Error</i>			
	Coefficiente	E. Estándar	Test t
Constantes específicas			
<b>Solo daños</b>	0	fijo	
<b>Heridos</b>	-0,86	0,08	-10,4
<b>Muertos</b>	-2,03	0,139	-14,59
<b>Características de la vía</b>			
<b>Heridos</b>			
No Hay Presencia de berma	0,135	0,057	2,36
Presencia de curva en 100 metros de vía	-0,184	0,052	-3,54
TPDA	-0,183	0,079	-2,31
señalización horizontal en buen estado	-0,079	0,058	-1,36
porcentaje de motos en la vía	0,267	0,162	1,65
presencia de peatones en 100 metros de vía	0,189	0,037	5,10
presencia de bicicletas en 100 metros de vía	0,192	0,044	4,35
pendiente mayor al 4%	-0,285	0,058	-4,96
<b>Muertos</b>			
<b>Características de la vía</b>			

<i>Mixto Con Componentes De Error</i>			
	<b>Coefficiente</b>	<b>E. Estándar</b>	<b>Test t</b>
Ancho de carril mayor 3.25 metros	0,320	0,110	2,91
No Hay Presencia de berma	0,305	0,089	3,43
presencia de curva en 100 metros de vía	-0,348	0,075	-4,64
TPDA	-0,435	0,115	-3,79
señalización horizontal en buen estado	-0,221	0,090	-2,47
porcentaje de motos en la vía	0,815	0,226	3,6
presencia de peatones en 100 metros de vía	0,114	0,052	2,17
presencia de bicicletas en 100 metros de vía	0,0538	0,063	0,86
pendiente mayor al 4%	-0,409	0,082	-5,02
$\sigma$	-0,426	-0,485	-0,88
<b>Parámetros</b>		20	
<b>Tamaño de la Muestra</b>		11271	
<b>Log-verosimilitud</b>		-10.087,3	

**Tabla 10 Modelo Logit Mixto Con Componentes De Error**  
**Fuente: Elaboración propia**

La Tabla 10 presenta los resultados obtenidos del modelo logit mixto con componentes de error. En la tabla se puede observar que la varianza del termino error no resultó ser significativa aunque arrojó que a medida que los anchos de carriles son mayores a 3.25 m y si no hay presencia de berma a la derecha del conductor, aumenta la probabilidad de que la gravedad del accidente sea más severo. Lo anterior puede estar correlacionado con la velocidad, ya que en esta vías el conductor puede percibir un nivel de confianza más alto que cuando transita por carriles angostos. Este nivel de confianza puede conllevar a que se presenten velocidades más altas lo que hace que al ocurrir un accidente este sea más severo.

## **5.2. Modelos Estimados De La Severidad De La Lesión De La Víctima Con La Base De Datos (2011 -2012)**

Utilizando la base de datos del periodo de (2011 – 2012) se estimó un modelo Logit Multinomial dado que se tienen sólo dos alternativas: muertos y heridos. Es de anotar que este modelo presenta un enfoque diferente a los modelos estimados para la base de datos (2007 – 2009), ya que este modelo se enfoca en determinar la severidad de la lesión de la víctima que pueda resultar en heridos o muertos.

La combinación lineal entre las variables y los coeficientes para describir la gravedad  $j$  de cada víctima  $q$  se presenta en las ecuaciones a continuación:

$$V_{Solodaños} = ASC1 * one \quad (37)$$

$$\begin{aligned}
V_{\text{Heridos}} = & ASC2 * one + B1_{\text{Berma}} * X1_{\text{Berma}} + B2_{\text{curva}} * X2_{\text{curva}} + B3_{\text{TPDA}} \\
& * X3_{\text{TPDA}} + B4_{\text{Sexo}} * X4_{\text{Sexo}} + B5_{\text{Edad}} * X5_{\text{Edad}} + B6_{\text{unión libre}} \\
& * X6_{\text{Unión libre}} + B7_{\text{Soltero}} * X7_{\text{Soltero}} + B8_{\text{Peatón}} * X8_{\text{Peatón}} \\
& + B9_{\text{Conductor}} * X9_{\text{Conductor}} + B10_{\text{Edu.primaria}} * X10_{\text{Edu.primaria}} \\
& + B11_{\text{Exc-alcohol}} * X11_{\text{Exc-alcohol}} + B13_{\text{Moto}} * X13_{\text{Moto}} \\
& + B14_{\text{Auto}} * X14_{\text{Auto}} + B15_{\text{Bus}} * X15_{\text{Bus}}
\end{aligned}
\quad (18)$$

<b>LOGIT MULTINOMIAL</b>			
	<b>Coefficiente</b>	<b>E. Estándar</b>	<b>Test t</b>
Constantes específicas			
<b>Heridos</b>	0	fijo	
<b>Características de la vía</b>			
<b>Heridos</b>			
No Hay Presencia de berma	0,402	0,146	2,75
Presencia de curva en 100 metros de vía	-0,39	0,149	-2,62
TPDA	-0,318	0,0740	-4.3
<b>Características de la Víctima</b>			
si la víctima es masculino	0,348	0,165	2.11
si la víctima tiene más de 40 años	-0,470	0.161	-2.91
si la víctima vive en unión libre	-1.05	0,198	-5,32
si la víctima es soltero	1,00	0,191	5,25
si la víctima es peatón	0,403	0,197	5,25
si la víctima es conductor	1,29	0,202	6,41
si la víctima solo tiene un nivel de educación básica	0,433	0,177	2,44
si va a exceso de alcohol	0,593	0,155	3,83
<b>Características del vehículo</b>			
si la víctima va en una moto	1.55	0,201	7,72
si la víctima va en auto	-1,64	0,211	7.75
si la víctima va en bus	-1.31	0,285	-4.59
<b>Parámetros</b>		15	
<b>Tamaño de la Muestra</b>		1226	
<b>Log-verosimilitud</b>		-572,285	

**Tabla 11 Modelo Logit Multinomial**  
**Fuente: Elaboración propia**

La Tabla 11 presenta los resultados obtenidos del modelo estimado. Los resultados presentan que los hombres tienen más probabilidad de sufrir lesiones más graves. Al considerar la influencia de la edad de la víctima en los accidentes se concluye que usuarios de edades mayores a 40 años, tienen menor probabilidad de sufrir una lesión más grave.

Otra variable interesante es si la víctima es soltera ya que este tipo de personas tiene más probabilidad de sufrir lesiones más graves que las víctimas que viven en unión libre.

También se determinó que si el accidente se produce por exceso de alcohol, existe mayor probabilidad de sufrir lesiones más graves, incluso con llevando a la muerte.

El modelo nos arroja que cuando la vía no presenta bermas, la severidad de la lesión de los accidentes aumenta. También se puede deducir que la severidad de la lesiones de los accidentes disminuye cuando hay presencia de curvas en 100 metros de vía. Lo anterior puede tener relación con la velocidad, ya que si hay presencia de curvas los usuarios tienden a disminuir su velocidad.

A medida que el flujo vehicular aumenta disminuye la severidad de la lesión en las víctimas involucradas en los accidentes. Esto debe tener cierta relación con la velocidad de los vehículos. A mayor flujo los vehículos tienden a ir más perceptivos y a bajas velocidades, lo que les permite estar pendientes de las condiciones de la vía. Lo contrario sucede en vías donde el flujo vehicular disminuye.

Teniendo en cuenta el tipo de vehículo en el que fue el accidente, se puede concluir que si la víctima va en una motocicleta tiene más probabilidad de tener una lesión más grave que si va en automóvil o en bus.



## 5 CONCLUSIONES.

La presente investigación se ha dedicado al estudio de determinar los factores de las características de la vía que intervienen en la severidad de los accidentes de tránsito. Para ello se ha utilizado modelos econométricos de la familia logit, los cuales permiten determinar de manera adecuada la severidad de los accidentes.

En la revisión bibliografía se expone y se concluye que los modelos de elección discreta cumplen los requerimientos para analizar los factores más determinantes en la severidad de un accidente de tránsito. Se expone que la teoría de utilidad aleatoria presenta una buena aproximación para analizar este tipo de datos.

Los resultados de la estimación de los modelos para explicar la gravedad de accidentes en carreteras colombianas muestran que los modelos ordinales fueron los que mejores resultados presentaron cuando se trabaja con datos a nivel de accidente y no de víctima.

Al realizar una comparación entre los modelos de componentes de error y el logit Multinomial se pudo concluir que la varianza del termino error no resultó ser significativa. La simplicidad del modelo logit multinomial es preferible en este caso cuando se compara frente al modelo de componentes de error.

Realizando un comparativo entre el modelo logit multinomial y el modelo logit ordinal se concluyó que el modelo que mejor se ajusta a los datos es el ordinal. El modelo logit ordinal tiene en cuenta la naturaleza de los datos, ya que este modelo reconoce que es más severo un accidente con muerto que un accidente con heridos y un accidente con heridos que un accidente con solo daño.

Igualmente se observó que la variable correspondiente a la velocidad de operación está altamente correlacionada con otras variables como los anchos de carril, la presencia de curvas en 100 metros de vía. Lo anterior porque en vías con anchos de carriles mayores a 3.0 metros los usuarios tiende a aumentar la velocidad, lo contrario se presenta cuando en 100 metros de vía se presenta curvas, donde el usuario disminuye la velocidad. Si se disminuye la velocidad el accidente será menos severo, mientras que cuando se experimente una alta velocidad más

aumenta la probabilidad de lesiones graves. A mayor velocidad, se incrementa la energía y se reducen las posibilidades de efectuar maniobras de emergencia. Este hecho invita a revisar los límites de velocidad teniendo en cuenta las condiciones de la infraestructura.

Otra variable relevante es el estado de la señalización horizontal, la cual aumenta la probabilidad de accidentes solo daño y disminuye los accidentes con lesiones graves.

El análisis estadístico de los modelos econométricos indica que el modelo logit anidado tiene un mejor poder explicativo que el modelo logit multinomial, pues presenta un mejor ajuste al considerar la correlación de los accidentes con consecuencias. El logit multinomial, a su vez, presenta un mejor ajuste que los modelos de componentes de error.

La seguridad en las carreteras colombianas requiere de una responsabilidad compartida. La reducción de la gravedad de los accidentes de tránsito solicita el compromiso y decisión del gobierno para adoptar leyes que castiguen la imprudencia de los usuarios, así como la participación de diferentes profesionales, como ingenieros viales que se preocupen más por la seguridad de las vías a diseñar.

Una vez concluida la tesis, se considera interesante investigar sobre otros aspectos relacionados con los factores que influyen en la severidad de los accidentes en las carreteras colombianas y se proponen.

- aproximaciones econométricas más complejas, involucrando parámetros aleatorios y variaciones sistemáticas en el efecto marginal de las variables consideradas.
- estimar modelos que involucren variables relacionadas con las características de los vehículos y de los individuos involucrados en el evento.
- Analizar modelos econométricos las cuales se basen en datos de accidentes viales en zonas urbanas colombianas.

## REFERENCIAS.

Peter t. Savolaine, Fred I. Manneing, Dominique lord, Mohammed A. Quddus 2011

Abdel-Aty et al., 1998 M. Abdel-Aty, C. Chen, J. Schott An Assessment of the Effect of Driver Age on Traffic Accident Involvement Using Log-linear Models Accident Analysis and Prevention, 30 (6) (1998), pp. 851–861

Eluru et al., 2008 N. Eluru, C. Bhat, D. Hensher A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes Accident Analysis and Prevention, 40 (3) (2008), pp. 1033–1054

Haque et al., 2010 M.M. Haque, H.C. Chin, H. Huang Applying Bayesian hierarchical models to examine motorcycle crashes at signalized intersections Accident Analysis & Prevention, 42 (2010), pp. 203–212

Kim et al., 2010 J.-K. Kim, G. Ulfarsson, V. Shankar, F. Mannering A note on modeling pedestrian injury severity in motor vehicle crashes with the mixed logit model Accident Analysis and Prevention, 42 (6) (2010), pp. 1751–1758

Malyshkina and Mannering, 2008 N. Malyshkina, F. Mannering Effect of increases in speed limits on severities of injuries in accidents Transportation Research Board, 2083 (2008), pp. 122–127

Milton et al., 2008 J. Milton, V. Shankar, F. Mannering Highway accident severities and the mixed logit model: an exploratory empirical analysis Accident Analysis and Prevention, 40 (1) (2008), pp. 260–266

Shankar and Mannering, 1996 V. Shankar, F. Mannering An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity Journal of Safety Research, 27 (3) (1996), pp. 183–194

Ulfarsson and Mannering, 2004 G. Ulfarsson, F. Mannering Differences in male and female injury severities in sport-utility vehicle, minivan, pickup and passenger car accidents Accident Analysis and Prevention, 36 (2) (2004), pp. 135–147

Washington et al., 2011 S.P. Washington, M.G. Karlaftis, F.L. Mannering Statistical and Econometric Methods for Transportation Data Analysis (second ed.) Chapman & Hall/CRC (2011)

Yang et al., 2011 Z. Yang, L. Zhibin, L. Pan, Z. Liteng Exploring contributing factors to crash injury severity at freeway diverge areas using ordered probit model Procedia Engineering, 21 (2011), pp. 178–185

Yamamoto et al., 2008 T. Yamamoto, J. Hashiji, V. Shankar Underreporting in traffic accident data, bias in parameters and the structure of injury severity models Accident Analysis and Prevention, 40 (4) (2008), pp. 1320–1329

F. Ye, D. Lord Comparing three commonly used crash severity models on sample size requirements: multinomial logit, ordered probit and mixed logit models Proceedings of the 90th Annual Meeting of the Transportation Research Board, Washington, DC (2011)

Plan Nacional de seguridad vial Colombia 2011- 2016Plan Mundial Para El Decenio De Acción Para La Seguridad Vial 2011-2020

Informe Sobre La Situación Mundial De La Seguridad Vial 2013

Comportamiento De Muertes Y Lesiones Por Accidente De Transporte, Colombia, 2013 Medicina Legal

Guía Técnica Para El Diseño De Las Zonas Laterales Para Vías Más Seguras; Fondo De Prevención Vial Colombia.

## LISTA DE ANEXOS.

ANEXO 1.	BASE DE DATOS PERIODO 2007 -2009	contiene un registro de 11.271 accidente
ANEXO 2.	BASE DE DATOS PERIODO 2011 -20012	contiene un registro de 1.226 accidente
ANEXO 3	MODELO MNL (2007-2009)	Contiene las especificaciones del modelo MNL
ANEXO 4	RESULTADOS MODELO MNL (2007-2009)	Contiene los resultados obtenidos para el modelo logit Multinomial escogido
ANEXO 5	MODELO OL (2007-2009)	Contiene las especificaciones del modelo logit ordinal
ANEXO 6	RESULTADOS MODELO OL (2007-2009)	Contiene los resultados obtenidos para el modelo logit ordinal escogido
ANEXO 7	MODELO LOGIT MIXTO CON COMPONENTES DE ERROR (2007-2009)	Contiene las especificaciones del modelo logit mixto con componentes de error
ANEXO 8	RESULTADOS LOGIT MIXTO CON COMPONENTES DE ERROR (2007-2009)	Contiene los resultados obtenidos para el modelo logit mixto con componentes de error escogidos
ANEXO 9	MODELO MNL (2011-2012)	Contiene las especificaciones del modelo logit Multinomial
ANEXO 10	RESULTADOS MODELO MNL (2011-2012)	Contiene los resultados obtenidos para el modelo logit Multinomial escogido